



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 201729

TO: Manjunath N Rao
Location: REM/2A01/2C70
Art Unit: 1652
September 19, 2006

Case Serial Number: 10/815774

From: P. Sheppard
Location: Remsen Building
Phone: (571) 272-2529

sheppard@uspto.gov

Search Notes

From: Chan, Christina
Sent: Thursday, September 14, 2006 3:56 PM
To: Rao, Manjunath N.; STIC-Biotech/ChemLib
Subject: RE: RUSH Sequence search for 10/815774

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Rao, Manjunath N.
Sent: Thursday, September 14, 2006 3:43 PM
To: Chan, Christina
Subject: RUSH Sequence search for 10/815774

Hello Christina

Please authorize the search request below as RUSH. The reason being this is a date goal case and claims are drawn to sequences.

Many Thanks
-Manjunath Rao

Manjunath N. Rao
Primary Examiner
Art Unit 1652
Ph: 571-272-0939

From: Manjunath N. Rao
Art Unit 1652, Room 2A01
Mail Box in Room 2C70
Phone: 272-0939

Date: 9-14-06

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

Please search the following as soon as possible for application with serial number
10/815,774

1. SEQ ID NO: 5 against all commercial nucleic acid sequence databases, issued patents/published applications nucleic acid sequence database and pending application nucleic acid sequence database. Please provide a print of results
2. SEQ ID NO: 6 against all commercial amino acid sequence databases, issued patents/published applications amino acid sequence database and pending application amino acid sequence database. Please provide a print of results.

If you have any questions please call me at the above phone number.

Thanks

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2006, 14:47:34 / Search time 3717 Seconds
(without alignments)
9548.243 Million cell updates/sec

Title: US-10-815-774-5

Perfect score: 555
Sequence: 1 ttggcgagtgtaattcctga.....atgtaccagctgtgattc 555

Scoring table: IDENTITY NJC
Gapop 10.0, Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl:*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vi:*
11: gb_ov:*
12: gb_hlg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	555	100.0	555	2	BD015599 Novel pro
2	555	100.0	555	2	AR560977 Sequence
3	555	100.0	555	2	AX113621 Sequence
4	555	100.0	1080	2	BD015600 Novel pro
5	555	100.0	1080	2	AR560978 Sequence
6	555	100.0	1080	2	AX113623 Sequence
7	555	100.0	1380	15	AB046594 Chryseoba
8	344.2	62.0	555	2	AR159957 Sequence
9	344.2	62.0	555	2	ES9337 Novel prote
10	344.2	62.0	555	2	AR570507 Sequence
11	344.2	62.0	1080	2	AR159962 Sequence
12	344.2	62.0	1080	2	ES9342 Novel prote
13	344.2	62.0	1080	2	AR570512 Sequence
14	43	7.7	239553	12	AC098659 Rattus no
15	41	7.4	125020	5	AF429315 Homo sapi
16	40.8	7.4	168540	12	AC129044 Rattus no
17	39.8	7.2	152202	5	AC094081 Homo sapi
18	39.8	7.2	231746	5	AC011389 Homo sapi

c	19	39.6	7.1	110000	12	CR385089_0	CR385089 Danto rer
c	20	39.6	7.1	110000	12	CR386985_1	Continuation (2 of
c	21	39.6	7.1	197752	12	CR387933	CR387933 Danto rer
c	22	39.6	7.1	264742	12	AC162403	AC162403 Bos tauru
c	23	39.6	7.1	239600	12	AC168301	AC168301 Bos tauru
c	24	38.8	7.0	154058	14	AC150680	AC150680 Monodelph
c	25	38.8	7.0	191715	12	AC178434	AC178434 Strongylo
c	26	38.6	7.0	155292	12	AC177955	AC177955 Strongylo
c	27	38.4	6.9	91275	13	AC152924	AC152924 Trybolium
c	28	38.4	6.9	192706	6	AC154283	AC154283 Mus muscu
c	29	38.2	6.9	202982	11	CR788249	CR788249 Zebrafish
c	30	38	6.8	27068	13	CEC15K7	Z67735 Caenorhabdi
c	31	38	6.8	110000	15	BX908798_00	BX908798 Parachlam
c	32	38	6.8	138862	11	BX908778	BX908778 Zebrafish
c	33	38	6.8	178077	11	AL929090	AL929090 Zebrafish
c	34	38	6.8	185947	5	AC092630	AC092630 Homo sapi
c	35	38	6.8	245721	12	AC172971	AC172971 Bos tauru
c	36	37.8	6.8	110000	12	BX897735_1	Continuation (2 of
c	37	37.8	6.8	110000	12	BX897735_2	Continuation (3 of
c	38	37.8	6.8	110017	12	CT573372	CT573372 Danto rer
c	39	37.8	6.8	147692	12	CT574555	CT574555 Danto rer
c	40	37.8	6.8	161731	11	AL590155	AL590155 Zebrafish
c	41	37.8	6.8	173576	12	AC166881	AC166881 Oryctolag
c	42	37.8	6.8	21971	6	AL954725	AL954725 Mus muscu
c	43	37.6	6.8	157430	6	AC135378	AC135378 Mus muscu
c	44	37.6	6.8	161223	11	AC022390	AC022390 Homo sapi
c	45	37.6	6.8	184556	6	AC117230	AC117230 Mus muscu

ALIGNMENTS

RESULT 1	BD015599	555 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD015599				
DEFINITION	Novel protein deamidase, microorganism producing the same, gene encoding the same, process for producing the same and use of the same.				
ACCESSION	BD015599				
VERSION	BD015599.1	GI:22556736			
KEYWORDS	JP 2001218590-A/3.				
SOURCE	Chryseobacterium sp. No. 9670				
ORGANISM	Chryseobacterium sp. No. 9670				
REFERENCE	1 (bases 1 to 555)				
AUTHORS	Yamaguchi, S.				
TITLE	Novel protein deamidase, microorganism producing the same, gene encoding the same, process for producing the same and use of the same.				
JOURNAL	Patent: JP 2001218590-A 3 14-AUG-2001;				
COMMENT	AMANO ENZYME CO LTD				
OS	Chryseobacterium sp. No. 9670				
PN	JP 2001218590-A/3				
PD	14-AUG-2001				
PF	04-DEC-2000 JP 2000368983				
PI	SHOTARO YAMAGUCHI				
PC	CI2N15/09:A23J3/34,C07K1/113,CI2N1/15,CI2N1/19,CI2N1/20,CI2N1/21,CI2N5/10,				
PC	CI2N9/80/(CI2N1/20,CI2R1/01),(CI2N9/80,CI2R1/01),CI2N15/00,				
PC	CI2N5/00				
CC	Novel protein deamidase, microorganism producing the same, CC				
Gene encoding					
CC	the same, process for producing the same and use of the same				
FT	Key				
FT	Location/Qualifiers				
FT	source				
FEATURES					
source	1..555				
	/organism="Chryseobacterium sp. No. 9670"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:161487"				
ORIGIN					

Query Match 100.0%; Score 555; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.1e-167;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGAGTGAATTCCTGATGTAGTACATTAATTCCTTAATTCATCAATAAAGAT 60
DB 1 TTGGCGAGTGAATTCCTGATGTAGTACATTAATTCCTTAATTCATCAATAAAGAT 60

QY 61 CAGCTTGGGGTACCTCTACAGGCGCTCTCACAGCATCATCAATTCCTGTAGAC 120
DB 61 CAGCTTGGGGTACCTCTACAGGCGCTCTCACAGCATCATCAATTCCTGTAGAC 120

QY 121 GGATGTTATGCAAGAGCCATTAAGATGAGCAATCTTAATGAACAAGCGGTATGACTGT 180
DB 121 GGATGTTATGCAAGAGCCATTAAGATGAGCAATCTTAATGAACAAGCGGTATGACTGT 180

QY 181 GAAATCAATTTGTATACGGAACCTTAAGGCATCAACAGAACTTGCTGTGGCGTGG 240
DB 181 GAAATCAATTTGTATACGGAACCTTAAGGCATCAACAGAACTTGCTGTGGCGTGG 240

QY 241 AGCTACACGTTGCAATTTGTAGTACATTAATAATGCTCCGAGTAAAGCAAAAAGA 300
DB 241 AGCTACACGTTGCAATTTGTAGTACATTAATAATGCTCCGAGTAAAGCAAAAAGA 300

QY 301 ATTATGATCCTTCACTATTTTCAAGCGGCTCTGTAAACAGATACAGATGAGAAACGCT 360
DB 301 ATTATGATCCTTCACTATTTTCAAGCGGCTCTGTAAACAGATACAGATGAGAAACGCT 360

QY 361 TGCCTTAACACCTTGGGATCTGCATCCGTTTCTCTTAATGCTAAATCTGCAGAAAT 420
DB 361 TGCCTTAACACCTTGGGATCTGCATCCGTTTCTCTTAATGCTAAATCTGCAGAAAT 420

QY 421 GTTATTACAGAACTCTAGTAATTTCTTACTGTATGACAAATCTGATCAATTAACCAAC 480
DB 421 GTTATTACAGAACTCTAGTAATTTCTTACTGTATGACAAATCTGATCAATTAACCAAC 480

QY 481 TGTGTACTGACTAAATTTTCACTGCTTCCGATGTTCTCTTCACTGCACCGGATGTA 540
DB 481 TGTGTACTGACTAAATTTTCACTGCTTCCGATGTTCTCTTCACTGCACCGGATGTA 540

QY 541 TCCAGCTGTGATTT 555
DB 541 TCCAGCTGTGATTT 555

RESULT 2
AR560977 AR560977 555 bp DNA linear PAT 08-OCT-2004
LOCUS Sequence 5 from patent US 6756221.
DEFINITION AR560977
ACCESSION AR560977
VERSION AR560977.1 GI:53973496
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 555)
AUTHORS Yamaguchi, S.
TITLE Protein-deamidating enzyme, microorganism producing the same, gene encoding the same, production process therefor, and use thereof
JOURNAL Patent: US 6756221-A 5 29-JUN-2004;
Amano Enzyme Inc.; Aichi;
JPX;

FEATURES
source Location/Qualifiers
1..555 /organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 555; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.1e-167;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGAGTGAATTCCTGATGTAGTACATTAATTCCTTAATTCATCAATAAAGAT 60

DB 1 TTGGCGAGTGAATTCCTGATGTAGTACATTAATTCCTTAATTCATCAATAAAGAT 60

QY 61 CAGCTTGGGGTACCTCTACAGGCGCTCTCACAGCATCATCAATTCCTGTAGAC 120
DB 61 CAGCTTGGGGTACCTCTACAGGCGCTCTCACAGCATCATCAATTCCTGTAGAC 120

QY 121 GGATGTTATGCAAGAGCCATTAAGATGAGCAATCTTAATGAACAAGCGGTATGACTGT 180
DB 121 GGATGTTATGCAAGAGCCATTAAGATGAGCAATCTTAATGAACAAGCGGTATGACTGT 180

QY 181 GAAATCAATTTGTATACGGAACCTTAAGGCATCAACAGAACTTGCTGTGGCGTGG 240
DB 181 GAAATCAATTTGTATACGGAACCTTAAGGCATCAACAGAACTTGCTGTGGCGTGG 240

QY 241 AGCTACACGTTGCAATTTGTAGTACATTAATAATGCTCCGAGTAAAGCAAAAAGA 300
DB 241 AGCTACACGTTGCAATTTGTAGTACATTAATAATGCTCCGAGTAAAGCAAAAAGA 300

QY 301 ATTATGATCCTTCACTATTTTCAAGCGGCTCTGTAAACAGATACAGATGAGAAACGCT 360
DB 301 ATTATGATCCTTCACTATTTTCAAGCGGCTCTGTAAACAGATACAGATGAGAAACGCT 360

QY 361 TGCCTTAACACCTTGGGATCTGCATCCGTTTCTCTTAATGCTAAATCTGCAGAAAT 420
DB 361 TGCCTTAACACCTTGGGATCTGCATCCGTTTCTCTTAATGCTAAATCTGCAGAAAT 420

QY 421 GTTATTACAGAACTCTAGTAATTTCTTACTGTATGACAAATCTGATCAATTAACCAAC 480
DB 421 GTTATTACAGAACTCTAGTAATTTCTTACTGTATGACAAATCTGATCAATTAACCAAC 480

QY 481 TGTGTACTGACTAAATTTTCACTGCTTCCGATGTTCTCTTCACTGCACCGGATGTA 540
DB 481 TGTGTACTGACTAAATTTTCACTGCTTCCGATGTTCTCTTCACTGCACCGGATGTA 540

QY 541 TCCAGCTGTGATTT 555
DB 541 TCCAGCTGTGATTT 555

RESULT 3
AX113621 AX113621 555 bp DNA linear PAT 30-MAY-2001
LOCUS Sequence 5 from Patent EP1106696.
DEFINITION AX113621
ACCESSION AX113621
VERSION AX113621.1 GI:13939799
KEYWORDS
SOURCE Chryseobacterium sp. No. 9670
ORGANISM Chryseobacterium sp. No. 9670
REFERENCE 1
AUTHORS Yamaguchi, S.
TITLE Protein-deamidating enzyme, microorganism producing the same, gene encoding the same, production process therefor, and use thereof
JOURNAL Patent: EP 1106696-A 5 13-JUN-2001;
Amano Enzyme Inc. (JP)

FEATURES
source Location/Qualifiers
1..555 /organism="Chryseobacterium sp. No. 9670"
/mol_type="unassigned DNA"
/db_xref="taxon:161487"

ORIGIN
Query Match 100.0%; Score 555; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.1e-167;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGAGTGAATTCCTGATGTAGTACATTAATTCCTTAATTCATCAATAAAGAT 60
DB 1 TTGGCGAGTGAATTCCTGATGTAGTACATTAATTCCTTAATTCATCAATAAAGAT 60

QY 61 CAGCTTGGGGTACCTCTACAGGCGCTCTCACAGCATCATCAATTCCTGTAGAC 120

Db 61 CAGCTCTGGGAGTACCTCTAGCGGCTCTCAACCATGATGACATTCAGATATCTGTAGAC 120
QY 121 GGATGTTATGCAAGAGCCCTAGATGAGCAATCTTATGAAACGGCTATGACTGT 180
Db 121 GGATGTTATGCAAGAGCCCTAGATGAGCAATCTTATGAAACGGCTATGACTGT 180
QY 181 GAAAAACAATTTGTATACGAAACCTAAAGCATCAACAGAACTTGTGTGTGGGTGG 240
Db 181 GAAAAACAATTTGTATACGAAACCTAAAGCATCAACAGAACTTGTGTGTGGGTGG 240
QY 241 AGCTACCAAGTTCATATTTGTAGCTATATAAATGCTTCCGAGTAACGAAAAAGA 300
Db 241 AGCTACCAAGTTCATATTTGTAGCTATATAAATGCTTCCGAGTAACGAAAAAGA 300
QY 301 ATTATGATGCTTCACTATTTTCAAGGGGCTGTACAGATACAGATGAGAAAGGCT 360
Db 301 ATTATGATGCTTCACTATTTTCAAGGGGCTGTACAGATACAGATGAGAAAGGCT 360
QY 361 TGGGTTAACAACCTCTTCCGAGTCTGATCCGTTTCTCTATGCTATATCTGAGGAAT 420
Db 361 TGGGTTAACAACCTCTTCCGAGTCTGATCCGTTTCTCTATGCTATATCTGAGGAAT 420
QY 421 GTTATTTACAGAAAGTCTAGTAATTTTACCTGTATGACAAACATTTGATCAATCAAC 480
Db 421 GTTATTTACAGAAAGTCTAGTAATTTTACCTGTATGACAAACATTTGATCAATCAAC 480
QY 481 TGTGTACTGACTTAATTTTCACTGCTTCCGAGTGTCTTCCCTCACTGACCGGATGTA 540
Db 481 TGTGTACTGACTTAATTTTCACTGCTTCCGAGTGTCTTCCCTCACTGACCGGATGTA 540
QY 541 TCCAGCTGTGATTT 555
Db 541 TCCAGCTGTGATTT 555

RESULT 4
BD015600 1080 bp DNA linear PAT 27-AUG-2002
LOCUS Novel protein deamidase, microorganism producing the same, gene
DEFINITION encoding the same, process for producing the same and use of the
same.

ACCESSION BD015600.1 GI:22556737
VERSION JP 2001218590-A/4
KEYWORDS Chryseobacterium sp. No. 9670
SOURCE Chryseobacterium sp. No. 9670
ORGANISM Bacteria; Bacteroidetes; Flavobacteriales;
Flavobacteriaceae; Chryseobacterium.

REFERENCE 1 (bases 1 to 1080)
Yamaguchi, S.
Novel protein deamidase, microorganism producing the same, gene
encoding the same, process for producing the same and use of the
Patent: JP 2001218590-A 4 14-AUG-2001;
JOURNAL AMANO ENZYME CO LTD
COMMENT OS Chryseobacterium sp. No. 9670
PN JP 2001218590-A/4
PD 14-AUG-2001
PF 04-DEC-2000 JP 2000368983
PI SHOTARO YAMAGUCHI
PC C12N15/09, A23J3/34, C07K1/13, C12N1/15, C12N1/19, C12N1/20, C12N1/
PC 21, C12N5/10,
PC C12N9/80//((C12N1/20, C12R1.01), (C12N9/80, C12R1.01), C12N15/00,
PC C12N5/00
CC Novel protein deamidase, microorganism producing the same, CC
gene encoding
the same, process for producing the same and use of the same

FEATURES
source 1. .1080
Location/Qualifiers
FT mat_peptide (466)..(1020).
FT CDS (61)..(1020)
FT Key location/Qualifiers
FT CDS (61)..(1020)
FT mat_peptide (466)..(1020).
FT Location/Qualifiers
1. .1080
/organism="Chryseobacterium sp. No. 9670"

ORIGIN
/mol_type="genomic DNA"
/db_xref="taxon:161487"

Query Match 100.0%; Score 555; DB 2; Length 1080;
Best Local Similarity 100.0%; Pred. No. 3.4e-167;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCAGTGTATTTCTGATGTAGTACATTAATTTCTTATTCATCAATTAAGAAT 60
Db 466 TTGGCAGTGTATTTCTGATGTAGTACATTAATTTCTTATTCATCAATTAAGAAT 525
QY 61 CAGCTCTGGGAGTACCTTAACGGGCTCTGACATGATCAATTCAGATATCTGTAGAC 120
Db 526 CAGCTCTGGGAGTACCTTAACGGGCTCTGACATGATCAATTCAGATATCTGTAGAC 585
QY 121 GGATGTTATGCAAGAGCCCTAGATGAGCAAACTTTATGAAACAGGGCTATGACTGT 180
Db 586 GGATGTTATGCAAGAGCCCTAGATGAGCAAACTTTATGAAACAGGGCTATGACTGT 645
QY 181 GAAAAACAATTTGTATACGAAACCTAAAGCATCAACAGAACTTGTGTGGGTGG 240
Db 646 GAAAAACAATTTGTATACGAAACCTAAAGCATCAACAGAACTTGTGTGGGTGG 705
QY 241 AGCTACCAAGTTCATATTTGTAGCTATATAAATGCTTCCGAGTAACGAAAAAGA 300
Db 706 AGCTACCAAGTTCATATTTGTAGCTATATAAATGCTTCCGAGTAACGAAAAAGA 765
QY 301 ATTATGATGCTTCACTATTTTCAAGGGGCTGTACAGATACAGATGAGAAAGGCT 360
Db 766 ATTATGATGCTTCACTATTTTCAAGGGGCTGTACAGATACAGATGAGAAAGGCT 825
QY 361 TGGGTTAACAACCTCTTCCGAGTCTGATCCGTTTCTCTATGCTATATCTGAGGAAT 420
Db 826 TGGGTTAACAACCTCTTCCGAGTCTGATCCGTTTCTCTATGCTATATCTGAGGAAT 885
QY 421 GTTATTTACAGAAAGTCTAGTAATTTTACCTGTATGACAAACATTTGATCAATCAAC 480
Db 886 GTTATTTACAGAAAGTCTAGTAATTTTACCTGTATGACAAACATTTGATCAATCAAC 945
QY 481 TGTGTACTGACTTAATTTTCACTGCTTCCGAGTGTCTTCCCTCACTGACCGGATGTA 540
Db 946 TGTGTACTGACTTAATTTTCACTGCTTCCGAGTGTCTTCCCTCACTGACCGGATGTA 1005
QY 541 TCCAGCTGTGATTT 555
Db 1006 TCCAGCTGTGATTT 1020

RESULT 5
AR560978 1080 bp DNA linear PAT 08-OCT-2004
LOCUS AR560978
DEFINITION Sequence 7 from patent US 6756221.
ACCESSION AR560978
VERSION AR560978.1 GI:53973497
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1080)
Yamaguchi, S.
Protein-deamidating enzyme, microorganism producing the same, gene
encoding the same, production process therefor, and use thereof
Patent: US 6756221-A 7 29-OCT-2004;
JOURNAL Amano Enzyme Inc.; Atch1;
JPX;

FEATURES
source 1. .1080
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 555; DB 2; Length 1080;

Best Local Similarity 100.0%; Pred. No. 3.4e-167;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TTGGCAGGTGTAATTCCTGATGTAAGTACATTAATTCCTTATTCATCAATAAAGAAAT 60
Db 466 TTGGCAGGTGTAATTCCTGATGTAAGTACATTAATTCCTTATTCATCAATAAAGAAAT 525
QY 61 CAGTCTTGCGGTACCTCTACGGGCTCCTCAACATGCATCATTCAATATTCCTGTAGAC 120
Db 526 CAGCTTGGCGGTACCTCTACGGGCTCCTCAACATGCATCATTCAATATTCCTGTAGAC 585
QY 121 GATGTTATGCAAGAGCCCAATAAGATGAGCAAAATCTTAATGAAACAAGGCTATGACTGT 180
Db 586 GATGTTATGCAAGAGCCCAATAAGATGAGCAAAATCTTAATGAAACAAGGCTATGACTGT 645
QY 181 GAAAAACAATTTGTATACGAAACCTTAAGGCATCAACAGAACTTGCTGTGGCGTGG 240
Db 646 GAAAAACAATTTGTATACGAAACCTTAAGGCATCAACAGAACTTGCTGTGGCGTGG 705
QY 241 AGCTACCAAGTGTGCAATTTGTGTAAGTATTAATAATGCTTCCGAGTAAAGGAAAAAGA 300
Db 706 AGCTACCAAGTGTGCAATTTGTGTAAGTATTAATAATGCTTCCGAGTAAAGGAAAAAGA 765
QY 301 ATTATGATCCTTCACATATTTTCAAGCGGTCCTGTAAAGATACAGATGAGGAAACGCT 360
Db 766 ATTATGATCCTTCACATATTTTCAAGCGGTCCTGTAAAGATACAGATGAGGAAACGCT 825
QY 361 TGGCTTAAACCTCTTGCGGATCTGCATCCGTTTCTCTTATGCTAAATCTGCAGAAAT 420
Db 826 TGGCTTAAACCTCTTGCGGATCTGCATCCGTTTCTCTTATGCTAAATCTGCAGAAAT 885
QY 421 GTTATTAACGAATCCTAGTAATCTTACCTGTATACAAACATCGATCAATACCAAC 480
Db 886 GTTATTAACGAATCCTAGTAATCTTACCTGTATACAAACATCGATCAATACCAAC 945
QY 481 TGTGTACTGACTAAATTTTCACTGCTTCCGATGTTCTCCTTCACTGCACCGGATGTA 540
Db 946 TGTGTACTGACTAAATTTTCACTGCTTCCGATGTTCTCCTTCACTGCACCGGATGTA 1005
QY 541 TCCAGCTGTGGATTT 555
Db 1006 TCCAGCTGTGGATTT 1020
```

RESULT 6
AX113623 1080 bp DNA linear PAT 30-MAY-2001
LOCUS AX113623
DEFINITION Sequence 7 from Patent EP1106696.
ACCESSION AX113623
VERSION AX113623.1 GI:13939800
KEYWORDS
SOURCE Chryseobacterium sp. No. 9670
ORGANISM Chryseobacterium sp. No. 9670
Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales;
Flavobacteriaceae; Chryseobacterium.

REFERENCE
AUTHORS Yamaguchi, S.
TITLE Protein-deamidating enzyme, microorganism producing the same, gene
encoding the same, production process therefor, and use thereof
JOURNAL Patent: EP 1106696-A 7 13-JUN-2001;
Amano Enzyme Inc. (JP)

FEATURES
source location/Qualifiers
1..1080 /organism="Chryseobacterium sp. No. 9670"
/mol_type="unassigned DNA"
/db_xref="taxon:161487"
61..1023
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC37857.1"
/db_xref="GI:13939801"
/translation="MKULFLSMMAFVTVLTFNSCADNSNGNDEINGEKLSVNSDKLKD

mat_peptide
/product="unnamed"

Query Match 100.0%; Score 555; DB 2; Length 1080;
Best Local Similarity 100.0%; Pred. No. 3.4e-167;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TTGGCAGGTGTAATTCCTGATGTAAGTACATTAATTCCTTATTCATCAATAAAGAAAT 60
Db 466 TTGGCAGGTGTAATTCCTGATGTAAGTACATTAATTCCTTATTCATCAATAAAGAAAT 525
QY 61 CAGTCTTGCGGTACCTCTACGGGCTCCTCAACATGCATCATTCAATATTCCTGTAGAC 120
Db 526 CAGTCTTGCGGTACCTCTACGGGCTCCTCAACATGCATCATTCAATATTCCTGTAGAC 585
QY 121 GATGTTATGCAAGAGCCCAATAAGATGAGCAAAATCTTAATGAAACAAGGCTATGACTGT 180
Db 586 GATGTTATGCAAGAGCCCAATAAGATGAGCAAAATCTTAATGAAACAAGGCTATGACTGT 645
QY 181 GAAAAACAATTTGTATACGAAACCTTAAGGCATCAACAGAACTTGCTGTGGCGTGG 240
Db 646 GAAAAACAATTTGTATACGAAACCTTAAGGCATCAACAGAACTTGCTGTGGCGTGG 705
QY 241 AGCTACCAAGTGTGCAATTTGTGTAAGTATTAATAATGCTTCCGAGTAAAGGAAAAAGA 300
Db 706 AGCTACCAAGTGTGCAATTTGTGTAAGTATTAATAATGCTTCCGAGTAAAGGAAAAAGA 765
QY 301 ATTATGATCCTTCACATATTTTCAAGCGGTCCTGTAAAGATACAGATGAGGAAACGCT 360
Db 766 ATTATGATCCTTCACATATTTTCAAGCGGTCCTGTAAAGATACAGATGAGGAAACGCT 825
QY 361 TGGCTTAAACCTCTTGCGGATCTGCATCCGTTTCTCTTATGCTAAATCTGCAGAAAT 420
Db 826 TGGCTTAAACCTCTTGCGGATCTGCATCCGTTTCTCTTATGCTAAATCTGCAGAAAT 885
QY 421 GTTATTAACGAATCCTAGTAATCTTACCTGTATACAAACATCGATCAATACCAAC 480
Db 886 GTTATTAACGAATCCTAGTAATCTTACCTGTATACAAACATCGATCAATACCAAC 945
QY 481 TGTGTACTGACTAAATTTTCACTGCTTCCGATGTTCTCCTTCACTGCACCGGATGTA 540
Db 946 TGTGTACTGACTAAATTTTCACTGCTTCCGATGTTCTCCTTCACTGCACCGGATGTA 1005
QY 541 TCCAGCTGTGGATTT 555
Db 1006 TCCAGCTGTGGATTT 1020
```

RESULT 7
AB046594 1380 bp DNA linear BCT 27-APR-2001
LOCUS AB046594
DEFINITION Chryseobacterium proteolyticum prga gene for protein-glutaminase,
complete cds.
ACCESSION AB046594
VERSION AB046594.1 GI:12597204
KEYWORDS
SOURCE Chryseobacterium proteolyticum
ORGANISM Chryseobacterium proteolyticum
Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales;
Flavobacteriaceae; Chryseobacterium.

REFERENCE
AUTHORS Yamaguchi, S., Jeenes, D.J., and Archer, D.B.
TITLE Protein-glutaminase from Chryseobacterium proteolyticum, an enzyme
that deamidates glutamyl residues in proteins. Purification,
characterization and gene cloning
JOURNAL Eur. J. Biochem. 268 (5), 1410-1421 (2001)
PUBMED 11231294

REFERENCE

2 (bases 1 to 1380)
Yamaguchi, S.
Direct Submission
Submitted (26-JUL-2000) Shofaro Yamaguchi, Amano Pharmaceutical Co.
Ltd., Research and Development, Suei, Kagamiyohara, Gifu 509-0108,
Japan (E-mail:LDV01447@nifty.ne.jp, tel:81-583-79-1220,
Fax:81-583-79-1232)

FEATURES

source

Location/Qualifiers

1. 1380

/organism="Chryseobacterium proteolyticum"

/mol_type="genomic DNA"

/strain="9670"

/db_xref="taxon:118127"

181. 1143

/gene="prga"

181. 1143

/gene="prga"

/codon_start=1

/transl_table=11

/product="protein-glutaminase"

/protein_id="BAB21508.1"

/db_xref="GI:12597205"

/translation="MKNLFLSMAPVTVTPNSCADSNGNGEINKEKISVNDKLD
FKTVPVGIIDENGMKVSFMLTAQVETETPTKENSQVIGMLQAVKESPVHFLKP
NSNEIGKVSASPEDVRYFKTLTKKVKQINKLASVDPVATLNSLFQIKNSCGT
STASPCITFRYVIDGCVARAHMRQILMANNQIDCKQFVYGNLKAISTGCVAMS
VALVSYKASGVTEKRIIDPSLFSGPTDTPAMRACVNTSGSASVSANTAGNV
YRSPNSNYLNNLNINVCVLTKEFLSLSGSPAPDVSSCGF"

181. 243

/gene="prga"

586. 1140

/gene="prga"

/product="unnamed"

/experiment="experimental evidence, no additional details
recorded"

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 555; DB 15; Length 1380;

Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTGGCAGTGTAAATTCCTGATAGTACATTAATCTTTATTCATCAATAAAGAAAT 60
|||||
586 TTGGCAGTGTAAATTCCTGATAGTACATTAATCTTTATTCATCAATAAAGAAAT 645
|||||
61 CAGTCTTGGGTAACCTCTAGGCGTCTCAACCATGATGATGATGATGATGATGATGAT 120
|||||
646 CAGTCTTGGGTAACCTCTAGGCGTCTCAACCATGATGATGATGATGATGATGATGAT 705
|||||
121 GGATGTTATGCAAGAGCCCAATAGATGAGCAAAATCTTAATGAACAAGCGCTATGACTGT 180
706 GGATGTTATGCAAGAGCCCAATAGATGAGCAAAATCTTAATGAACAAGCGCTATGACTGT 765
|||||
181 GAAAAACAATTTGTATACGAAACCTAAAGCATCAACAGAACTTGCTGTGTGCGTGTG 240
766 GAAAAACAATTTGTATACGAAACCTAAAGCATCAACAGAACTTGCTGTGTGCGTGTG 825
|||||
241 AGCTACCAAGTTCATATTGTGTAAGCTATAAAAATGCTTCCGAGTAAAGAAAAAGA 300
|||||
826 AGCTACCAAGTTCATATTGTGTAAGCTATAAAAATGCTTCCGAGTAAAGAAAAAGA 885
|||||
301 ATTATGATCCTTCACTATTTTCAAGCGGCTCGTAAAGATACAGATGAGAGAAAGCGT 360
886 ATTATGATCCTTCACTATTTTCAAGCGGCTCGTAAAGATACAGATGAGAGAAAGCGT 945
|||||
361 TGGGTTAACAACCTCTTGGGATCTGATCGCTTCTCTTATGCTAATCTGACGAGAAAT 420
946 TGGGTTAACAACCTCTTGGGATCTGATCGCTTCTCTTATGCTAATCTGACGAGAAAT 1005
|||||
421 GTTATTATTAAGAAGTCTTGAATTTCTTACCTGTATGACAAACATCTGATCAATCAAC 480
|||||
1006 GTTATTATTAAGAAGTCTTGAATTTCTTACCTGTATGACAAACATCTGATCAATCAAC 1065
|||||

QY 481 TGTGTACTGACTTAATTTTCACTGCTTCGGATGTTCTTCTTACCTGACCGGATGTA 540
|||||
DB 1066 TGTGTACTGACTTAATTTTCACTGCTTCGGATGTTCTTCTTACCTGACCGGATGTA 1125
|||||
QY 541 TCCAGCTGTGATTT 555
|||||
DB 1126 TCCAGCTGTGATTT 1140
|||||

RESULT 8

AR159957

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity 62.0%; Score 344.2; DB 2; Length 555;

Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

7 AGTGTAAATTCCTGATAGTACATTAATCTTTATTCATCAATAAAGAAATGCTCT 66
|||||
DB 7 AGTGTAAATTCCTGATAGTACATTAATCTTTATTCATCAATAAAGAAATGCTCT 66
|||||
67 TGGGTAACCTCTAGGCGTCTCAACCATGATGATGATGATGATGATGATGATGATGAT 126
|||||
67 TGGGTAACCTCTAGGCGTCTCAACCATGATGATGATGATGATGATGATGATGATGAT 126
|||||
127 TATGCAAGAGCCCAATAGATGAGCAAAATCTTAATGAACAAGCGCTATGACTGTGAAAA 186
127 TATGCAAGAGCCCAATAGATGAGCAAAATCTTAATGAACAAGCGCGGCTATGACTGTGAAAA 186
|||||
187 CAATTTGTATACGAAACCTAAAGCATCAACAGAACTTGCTGTGTGCGTGTGAGCTTAC 246
|||||
187 CAGTCTTGTATAGTAAATCTGAGAGCTTCTAAGAAACATGCTGTATCATGCGGTATAT 246
|||||
247 CAGTCTTGTATAGTAAATCTGAGAGCTTCTAAGAAACATGCTGTATCATGCGGTATAT 306
247 CAGTCTTGTATAGTAAATCTGAGAGCTTCTAAGAAACATGCTGTATCATGCGGTATAT 306
|||||
307 GATCTTCACTATTTTCAAGCGGCTCGTAAAGATACAGATGAGAGAAAGCGTGTG 366
307 GATCTTCACTATTTTCAAGCGGCTCGTAAAGATACAGATGAGAGAAAGCGTGTG 366
|||||
367 AACACCTCTTGGGATCTGATCGCTTCTTATGCTAATCTGACGAGAAATGTTAT 426
367 AACACCTCTTGGGATCTGATCGCTTCTTATGCTAATCTGACGAGAAATGTTAT 426
|||||
427 TACAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 486
427 TACAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 486
|||||
487 CTGACCTAATTTTCACTGCTTCGGATGTTCTTCTTACCTGACCGGATGATGATGAT 546
487 CTGACCTAATTTTCACTGCTTCGGATGTTCTTCTTACCTGACCGGATGATGATGAT 546
|||||
QY 547 TGTGATTT 555
|||||
DB 547 TGTGATTT 555
|||||

RESULT 9
E59337 555 bp DNA linear PAT 18-JUN-2001
LOCUS Novel protein deamidation enzyme, gene encoding it, process for
DEFINITION producing the same, and utilization thereof.
ACCESSION E59337
VERSION E59337.1 GI:13023304
KEYWORDS JP 2000050887-A/3.
SOURCE Chryseobacterium gleum
ORGANISM Chryseobacterium gleum
Bacteria; Bacteroidetes; Flavobacteriales;
Flavobacteriaceae; Chryseobacterium.
REFERENCE 1 (bases 1 to 555)
AUTHORS Shotaro, Y. and Akira, M.
TITLE Novel protein deamidation enzyme, gene encoding it, process for
producing the same, and utilization thereof
JOURNAL Patent: JP 2000050887-A 3 22-FEB-2000;
AMANO PHARMACEUT CO LTD
COMMENT OS Chryseobacterium gleum
PN JP 2000050887-A/3
PD 22-FEB-2000
PR 04-JUN-1999 JP 1999158703
PT SHOTARO YAMAGUCHI, AKIRA MATSUIRA
PC C12N15/09, C12N9/80//A21D2/36, A23C9/154, A23C11/06, A23J3/00, PC
A23J3/10,
A23J3/16, A23J3/18, A23J3/34, A23J1/176, A23J1/23, A23J1/317, PC
PC A23J3/80, C12R1:01, (C12N9/80, C12R1:20), C12N15/00 CC
FH Key Location/Qualifiers
FT source 1..555
FT Location/Qualifiers
1..555
/organism="Chryseobacterium gleum"
/mol_type="genomic DNA"
/db_xref="taxon:250"

ORIGIN
Query Match 62.0%; Score 344.2; DB 2; Length 555;
Best Local Similarity 76.7%; Pred. No. 3.2e-99;
Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 7 AGGTAAATTCCTGATGAGTACATTAATCTTTATTCATCAATAAAGAACTAGT 66
Db 7 AGGTAAATTCCTGATGAGTACATTAATCTTTATTCATCAATAAAGAACTAGT 66
QY 67 TGGGTAACCTCTAAGCGGCTCTCAACATGCATCAATTCCTGTAGACGATGT 126
Db 67 TGGGTAACCTCTAAGCGGCTCTCAACATGCATCAATTCCTGTAGACGATGT 126
QY 127 TATCAAGAGCCATAAGATGAGACAAATCTTAATGAACAAAGGCTATGACTGAAAA 186
Db 127 TATCAAGAGGCTCAAAAATGAGACAAATCTTAATGAACAAAGGCTATGACTGAAAA 186
QY 187 CAATTTGTATACGAAACCTTAAGCATCAACAGAAATTCGTGTGGGTGAGCTAC 246
Db 187 CAATTTGTATACGAAACCTTAAGCATCAACAGAAATTCGTGTGGGTGAGCTAC 246
QY 187 CAGTTCTATATGTGAATCTGAGAGCTTCTACAGAAACATCTGTATCATGGGTATAT 246
Db 187 CAGTTCTATATGTGAATCTGAGAGCTTCTACAGAAACATCTGTATCATGGGTATAT 246
QY 247 CAGGTGCAATATGTGTAAGCTATATAAATGCTTCGAGTAAAGGAAAAAGATTTT 306
Db 247 CAGGTGCAATATGTGTAAGCTATATAAATGCTTCGAGTAAAGGAAAAAGATTTT 306
QY 307 GATCCTTCAATATTTTCAAGCGGCTCTGTAACAGATACAGATGAGAAACGCTTGCTT 366
Db 307 GATCCTTCAATATTTTCAAGCGGCTCTGTAACAGATACAGATGAGAAACGCTTGCTT 366
QY 367 AACACCTTTGGGATCTGCATCCGTTTCTTATGCTAATCTGCGAGAAATGTTAT 426
Db 367 AACACCTTTGGGATCTGCATCCGTTTCTTATGCTAATCTGCGAGAAATGTTAT 426
QY 427 TACGAAGTCCGTAGTAATTTCTTACTGTATGACAAATCTGATCAATACAACTGTGTA 486
Db 427 TACGAAGTCCGTAGTAATTTCTTACTGTATGACAAATCTGATCAATACAACTGTGTA 486

Db 427 TACGAAGTCCGTAGTAATTTCTTACTGTATGACAAATCTGATCAATACAACTGTGTA 486
QY 487 CTGACCTAAATTTTCACTGCTTCGGAATGCTTCCTTCAACCGACCGGATGATCCAG 546
Db 487 TTAACATATTTCTCATCCCTTCAGAGATGTTCTTCCCGACGACCAAGTGTGCAAGC 546
QY 547 TGTGATTT 555
Db 547 TGTGATTT 555

RESULT 10
AR570507 555 bp DNA linear PAT 14-DEC-2004
LOCUS Sequence 5 from patent US 6770469.
DEFINITION AR570507
ACCESSION AR570507
VERSION AR570507.1 GI:56571305
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 555)
AUTHORS Yamaguchi, S. and Matsunura, A.
TITLE Protein-deamidating enzyme, gene encoding the same, production
process thereof, and use thereof
JOURNAL Patent: US 6770469-A 5 03-AUG-2004;
Amano Pharmaceutical Co., Ltd.; Atchii;
JPX;

FEATURES
source Location/Qualifiers
1..555
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 62.0%; Score 344.2; DB 2; Length 555;
Best Local Similarity 76.7%; Pred. No. 3.2e-99;
Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 7 AGGTAAATTCCTGATGAGTACATTAATCTTTATTCATCAATAAAGAACTAGT 66
Db 7 AGGTAAATTCCTGATGAGTACATTAATCTTTATTCATCAATAAAGAACTAGT 66
QY 67 TGGGTAACCTCTAAGCGGCTCTCAACATGCATCAATTCCTGTAGACGATGT 126
Db 67 TGGGTAACCTCTAAGCGGCTCTCAACATGCATCAATTCCTGTAGACGATGT 126
QY 127 TATCAAGAGCCATAAGATGAGACAAATCTTAATGAACAAAGGCTATGACTGAAAA 186
Db 127 TATCAAGAGGCTCAAAAATGAGACAAATCTTAATGAACAAAGGCTATGACTGAAAA 186
QY 187 CAATTTGTATACGAAACCTTAAGCATCAACAGAAATTCGTGTGGGTGAGCTAC 246
Db 187 CAATTTGTATACGAAACCTTAAGCATCAACAGAAATTCGTGTGGGTGAGCTAC 246
QY 187 CAGTTCTATATGTGAATCTGAGAGCTTCTACAGAAATCTGTATCATGGGTATAT 246
Db 187 CAGTTCTATATGTGAATCTGAGAGCTTCTACAGAAATCTGTATCATGGGTATAT 246
QY 247 CAGGTGCAATATGTGTAAGCTATATAAATGCTTCGAGTAAAGGAAAAAGATTTT 306
Db 247 CAGGTGCAATATGTGTAAGCTATATAAATGCTTCGAGTAAAGGAAAAAGATTTT 306
QY 307 GATCCTTCAATATTTTCAAGCGGCTCTGTAACAGATACAGATGAGAAACGCTTGCTT 366
Db 307 GATCCTTCAATATTTTCAAGCGGCTCTGTAACAGATACAGATGAGAAACGCTTGCTT 366
QY 367 AACACCTTTGGGATCTGCATCCGTTTCTTATGCTAATCTGCGAGAAATGTTAT 426
Db 367 AACACCTTTGGGATCTGCATCCGTTTCTTATGCTAATCTGCGAGAAATGTTAT 426
QY 427 TACGAAGTCCGTAGTAATTTCTTACTGTATGACAAATCTGATCAATACAACTGTGTA 486
Db 427 TACGAAGTCCGTAGTAATTTCTTACTGTATGACAAATCTGATCAATACAACTGTGTA 486
QY 487 CTGACCTAAATTTTCACTGCTTCGGAATGCTTCCTTCAACCGACCGGATGATCCAG 546
Db 487 CTGACCTAAATTTTCACTGCTTCGGAATGCTTCCTTCAACCGACCGGATGATCCAG 546

Db 487 TTAACATATTCATCCCTTCAGAGTGTCTCTCCACGACCAAGTGAAGAC 546
QY 547 TGTGATTT 555
Db 547 TGTGATTT 555

RESULT 11
LOCUS AR159962 1080 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 11 from patent US 6251651.
ACCESSION AR159962
VERSION AR159962.1 GI:16222841
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1080)
AUTHORS Yamaguchi, S. and Matsunura, A.
TITLE Protein-deamidating enzyme, gene encoding the same, production process thereof, and use thereof
JOURNAL Patent: US 6251651-A 11 26-JUN-2001;
FEATURES
source location/Qualifiers
1.1080
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 62.0%; Score 344.2; DB 2; Length 1080;
Best Local Similarity 76.7%; Pred. No. 3.5e-99;
Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 7 AGTGTATTCCTGATGAGTACATTAATCTTTATTCATGAATAAGATCAGTCT 66
Db 469 AGTGTATTCCTGATGAGTACATTAATCTTTATTCATGAATAAGATCAGTCT 528
QY 67 TGGGTAACCTCTACGGCGTCTCAGCATGATCAGATTCAGATTCCTGATGAGTAT 126
Db 529 TGGGTAACCTCTACGGCGTCTCAGCATGATCAGATTCCTGATGAGTAT 588
QY 127 TATGCAAGAGCCATTAAGATGAGACAAATCTTAATGAACAGCGCTATGACTGAAAA 186
Db 589 TATGCAAGAGCCATTAAGATGAGACAAATCTTAATGAACAGCGCTATGACTGAAAA 648
QY 187 CAATTTGTATAGGAAACCTTAAGGATCAACAGAACTTGTGTGCGGTGAGACTAC 246
Db 649 CAATTTGTATAGGAAACCTTAAGGATCAACAGAACTTGTGTGCGGTGAGACTAC 708
QY 247 CAGGTGCAATATTTGTAAGCTATTAATAATGCTTCGAGTAAAGAAAAAATTAAT 306
Db 709 CAGGTGCAATATTTGTAAGCTATTAATAATGCTTCGAGTAAAGAAAAAATTAAT 768
QY 307 GATCTTCACTATTTTCAAGCGGTCTGTAAAGATACAGATGAGAAACGCTTGCGTT 366
Db 769 GATCTTCACTATTTTCAAGCGGTCTGTAAAGATACAGATGAGAAACGCTTGCGTT 828
QY 367 AACACCTCTTGGCGATCTGATCCGTTCTCTTAATGCTAATCTGCAAGAAATGTTAT 426
Db 829 AACACCTCTTGGCGATCTGATCCGTTCTCTTAATGCTAATCTGCAAGAAATGTTAT 888
QY 427 TACAGAAGTCTAGTATTTCTTACCTGTATGACAAATCTGATCAATACCACTGTGA 486
Db 889 TACAGAAGTCTAGTATTTCTTACCTGTATGACAAATCTGATCAATACCACTGTGA 948
QY 487 CTGACTAAATTTTCACTGCTTTCGAGATGTTCTCTTCACTGACCGAGATGATTCAC 546
Db 949 TTAACATATTCATCCCTTCAGAGTGTCTCTCCACGACCAAGTGAAGAC 1008
QY 547 TGTGATTT 555
Db 1009 TGTGATTT 1017

RESULT 12
LOCUS E59342 1080 bp DNA linear PAT 18-JUN-2001
DEFINITION Novel protein deamidation enzyme, gene encoding it, process for producing the same, and utilization thereof.
ACCESSION E59342
VERSION E59342.1 GI:13023309
KEYWORDS JP 2000050887-A/8.
SOURCE Chryseobacterium gleum
ORGANISM Chryseobacterium gleum

REFERENCE 1 (bases 1 to 1080)
AUTHORS Shotaro, Y. and Akira, M.
TITLE Novel protein deamidation enzyme, gene encoding it, process for producing the same, and utilization thereof
JOURNAL Patent: JP 2000050887-A 8 22-FEB-2000;
COMMENT AMANO PHARMACEUT CO LTD
OS Chryseobacterium gleum
PN JP 2000050887-A/8
PD 22-FEB-2000
PR 04-JUN-1999 JP 1999158703

PC SHOTARO YAMAGUCHI, AKIRA MATSUNURA
PC C12N15/09, C12N9/80//A21D2/36, A23C9/154, A23C11/06, A23J3/00, PC
A23J3/10,
PC A23J3/16, A23J3/18, A23J3/34, A23L1/176, A23L1/23, A23L1/317, PC
(C12N9/80, C12N1/01), (C12N9/80, C12N1/20), C12N15/00 CC
FH Key Location/Qualifiers
FT CDS (61)..(1020)
FT mat peptide (463)..(1017).

FEATURES
source location/Qualifiers
1.1080
/organism="Chryseobacterium gleum"
/mol_type="genomic DNA"
/db_xref="taxon:250"

ORIGIN

Query Match 62.0%; Score 344.2; DB 2; Length 1080;
Best Local Similarity 76.7%; Pred. No. 3.5e-99;
Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 7 AGTGTATTCCTGATGAGTACATTAATCTTTATTCATGAATAAGATCAGTCT 66
Db 469 AGTGTATTCCTGATGAGTACATTAATCTTTATTCATGAATAAGATCAGTCT 528
QY 67 TGGGTAACCTCTACGGCGTCTCAGCATGATCAGATTCAGATTCCTGATGAGTAT 126
Db 529 TGGGTAACCTCTACGGCGTCTCAGCATGATCAGATTCCTGATGAGTAT 588
QY 127 TATGCAAGAGCCATTAAGATGAGACAAATCTTAATGAACAGCGCTATGACTGAAAA 186
Db 589 TATGCAAGAGCCATTAAGATGAGACAAATCTTAATGAACAGCGCTATGACTGAAAA 648
QY 187 CAATTTGTATAGGAAACCTTAAGGATCAACAGAACTTGTGTGCGGTGAGACTAC 246
Db 649 CAATTTGTATAGGAAACCTTAAGGATCAACAGAACTTGTGTGCGGTGAGACTAC 708
QY 247 CAGGTGCAATATTTGTAAGCTATTAATAATGCTTCGAGTAAAGAAAAAATTAAT 306
Db 709 CAGGTGCAATATTTGTAAGCTATTAATAATGCTTCGAGTAAAGAAAAAATTAAT 768
QY 307 GATCTTCACTATTTTCAAGCGGTCTGTAAAGATACAGATGAGAAACGCTTGCGTT 366
Db 769 GATCTTCACTATTTTCAAGCGGTCTGTAAAGATACAGATGAGAAACGCTTGCGTT 828
QY 367 AACACCTCTTGGCGATCTGATCCGTTCTCTTAATGCTAATCTGCAAGAAATGTTAT 426
Db 829 AACACCTCTTGGCGATCTGATCCGTTCTCTTAATGCTAATCTGCAAGAAATGTTAT 888
QY 427 TACAGAAGTCTAGTATTTCTTACCTGTATGACAAATCTGATCAATACCACTGTGA 486
Db 889 TACAGAAGTCTAGTATTTCTTACCTGTATGACAAATCTGATCAATACCACTGTGA 948

QY	487	CGA	CTAAATTTG	CTG	CTTCCG	GAAGTTCT	CTCTG	ACCG	GCACG	AGATTC	CAGC	546
Db	949	TTAA	CAATTTCT	CACTCC	CTTCAG	AGATTTCT	CTCTCC	CA	GCACCA	GAAGT	AGCAAGC	1008
QY	547	TGT	GATTT	555								
Db	1009	TGT	GATTT	1017								
RESULT 13												
AR570512		1080 bp		DNA		linear		PAT 14-DEC-2004				
LOCUS		Sequence 11 from patent US 6770469.										
DEFINITION		AR570512										
ACCESSION		AR570512										
VERSION		AR570512.1 GI:56571310										
KEYWORDS												
SOURCE		Unknown.										
ORGANISM		Unknown.										
REFERENCE		1 (bases 1 to 1080)										
AUTHORS		Yamaguchi,S. and Matsunura,A.										
TITLE		Protein-deamidating enzyme, gene encoding the same, production process thereof, and use thereof										
JOURNAL		Patent: US 6770469-A 11 03-AUG-2004;										
		Amano Pharmaceutical Co., Ltd.; Aichi, JPY;										
FEATURES												
source		Location/Qualifiers										
ORIGIN		1..1080										
		/organism="unknown"										
		/mol_type="genomic DNA"										
Query Match		62.0%; Score 344.2; DB 2; Length 1080;										
Best Local Similarity		76.7%; Pred. No.3.5e-99;										
Matches		421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;										
QY	7	AGT	GAATTC	CGATG	TAC	TAC	TAAATTC	TTTAT	TCAT	CAATAA	AGATC	66
Db	469	AGT	GTATTC	CGATG	TGCG	CAACG	CTGAAC	AGTTT	ATTTAC	CCCAAT	CAAAAAC	528
QY	67	TGG	GCATCT	CTAC	CGGCGT	CCTC	CACCAT	GCAT	CACAT	TTCAG	ATTC	126
Db	529	TGG	GGAATCT	TAC	AGCAT	CTTCTC	CTGTAT	CACCTT	CAGATAT	CCGGTGA	CGGATG	588
QY	127	TAT	GCAAG	CCCAT	TAAAT	TGAG	CAAAATCT	TAA	TGAACA	AGGCT	TATGAC	186
Db	589	TAT	GCAAGG	CCCTCA	AAAAT	TGAG	CAAAATCT	TAA	TGAAGCG	CGCTAT	GATGTA	648
QY	187	CAAT	TGTAT	GGA	AACTTAA	AGCAT	CAACAG	GAATTC	GTGTGT	GGCGTGA	CTAC	246
Db	649	CAG	TGTAT	GTAT	GTAT	GTAG	AGCTTCTA	CA	GGAACAT	GTGTGTAT	CAATGG	708
QY	247	CAG	TGCA	TAT	TGGTAA	GTAT	TAAAAAT	TGCTTC	CGAGTAA	CGAAAAA	GAATTA	306
Db	709	CAG	TGCAAT	TTTGGTAA	GTCTCA	AAATAT	TGCTTC	CAG	AATTTG	TGA	AAAAA	768
QY	307	GAT	CTTCA	CTATTT	CAAG	CGGTCT	CTG	TAA	CGAATG	ACGATG	AGAAAG	366
Db	769	GAT	CTTCAAT	TTTCTC	ACGCGGTCT	CTG	TAA	CGAATTC	GCAATG	AGAAAG	CTCAT	828
QY	367	AAC	ACCTTT	GCGAGAT	CTG	GCAT	CGTTTCT	CTTAT	ATG	CAATAC	TGCAG	426
Db	829	AA	ACCAAG	CTGGCG	AGATCT	GTGTAT	CTTCTT	CTA	AGCAAT	ACGAG	AAATGT	888
QY	427	TAC	GAAAGT	CTAGT	ATTTCTT	AC	CTGTAT	TGAC	ACA	CAATCT	GATCAAT	486
Db	889	TAC	GAAAGT	CTC	GAGTTCAT	TACT	GTATG	ATTAACA	ATAT	GTGAAT	TCAAT	948
QY	487	CTG	ACT	TAATTT	CAC	TGTTTCCG	ATGTTC	CTTAC	CGAC	CGAATG	ATTC	546
Db	949	TTTAA	CAATATTC	ATCCCTTT	CAG	AGATGTTCT	CTTCC	CCAC	GA	CAAGT	ATGAC	1006

OR	547	TGTCGATT	555	
LOCUS	AC098659			
DEFINITION	Rattus norvegicus clone CH230-100L14, WORKING DRAFT SEQUENCE, 4			
ACCESSION	AC098659	239553 bp	DNA linear HTG 10-MAY-2003	
VERSION	AC098659.7	GI:30520678		
KEYWORDS	HTG, HTGS, PHASE1, HTGS, DRAFT, HTGS, FULLTOP.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.			
REFERENCE	Murphy, D., Mearns, M., Lee, A., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Aryalpechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bielow, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Cencer, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Dreger, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falle, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, A., Foster, M., Foster, P., Fraser, C., Gabisi, A., Gall, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Giller, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, J., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louised, H., Lorado, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Manning, S., McLeod, M., McNeill, T., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeme, O., Okunnu, G., Olariunpangon, A., Pal, S., Parke, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, B., Pu, L., Pu, M., Quirroz, J., Rachlin, E., Reeves, K., Reiger, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsberg, A., Sison, I., Sitter, C., Sma, J., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Soe, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Vals, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wiczyk, R., Widen, H., Wootley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von Niederhausern, A., Weiss, R., Smith, D., Holt, R., Smith, H., O., Weinstein, G., and Gibbs, R. A.			
TITLE	Unpublished			
JOURNAL	2 (bases 1 to 239553)			
REFERENCE	Worley, K. C.			
AUTHORS	Direct Submission			
TITLE				

JOURNAL Submitted (28-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 239553)
REFERENCE Rat Genome Sequencing Consortium.
AUTHORS Direct Submission
TITLE Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL On May 10, 2003 this sequence version replaced gi:25091888.
COMMENT The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GHVY
Center clone name: CH230-100L14
----- Summary Statistics
Assembly program: Atlas 3.0:
Consensus quality: 223709 bases at least Q40
Consensus quality: 226612 bases at least Q20
Estimated insert size: 228176; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 230323: contig of 230323 bp in length
* 230324 230423: gap of unknown length
* 230424 231606: contig of 1183 bp in length
* 231607 231706: gap of unknown length
* 231707 233074: contig of 1368 bp in length
* 233075 233174: gap of unknown length
* 233175 239553: contig of 6379 bp in length.

FEATURES
source
1..239553
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-100L14"
1..1813
/note="wgs contig"
1157..12737
/note="wgs contig"
230324..230423
/estimated_length=unknown
231607..231706
/estimated_length=unknown
233075..233174
/estimated_length=unknown

misc_feature
misc_feature
gap
gap
gap
ORIGIN

Query Match 7.7%; Score 43; DB 12; Length 239553;
Best Local Similarity 60.9%; Pred. No. 0.12;
Matches 70; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 316 CTAATTTCAAGCGGCTCTTAACAGATACAGATGAGAAAGCTTGCTTAACCTCT 375
|||
Db 61050 CTCGCTCACTCTGTCATTTAAACAAAGAACTTCATCACTCACTGAGTGAACCTCAAT 61109
|||
QY 376 TCGGATTCGACGCGTTTCTCTTATGCTAATGACGAAATGTTATTACA 430
|||
Db 61110 AATGCTCTGTACCGTTTCTGTTATGAGAAAGAACTGAAATGTCATTTACA 61164
|||

RESULT 15
AF429315/c 125020 bp DNA linear PRI 18-JAN-2002
LOCUS AF429315
DEFINITION Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 125020)
Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
11694876
2 (bases 1 to 125020)
Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
Direct Submission
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA

FEATURES
source
1..125020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/note="Isolated from a patient with Huntington's
disease-Like 2 (HDL2)"
complement(35581..35746)
/rpt_type=tandem
/rpt_unit_seq="ctg"
complement(<36507..>36887)
/gene="JPH3"
/note="synonym: JP3"
complement(<36507..>36887)
/gene="JPH3"
/product="junctophilin 3"
/product="junctophilin 3"
complement(<36507..>36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AA140941.1"
/db_xref="GI:17646245"
/translation="MSGGRRFVFDGSGYCGGEMGKAKHGVCCTGPKQGGYTSWS
HGFEVLGYTPSGNTVGTMAQGRHGLGSKKWKWYKGMTGFKGRYVRECAQ
NGAKYEGTWSNGLQDGYETYSDC"

repeat_region
gene
mRNA
CDS
ORIGIN

Query Match 7.4%; Score 41; DB 5; Length 125020;
Best Local Similarity 12.8%; Pred. No. 0.5; Mismatches 154; Indels 0; Gaps 0;
Matches 47; Conservative 154; Mismatches 165; Indels 0;

```
QY      130  GCAAGAGCCATAGATGAGACAATCTTATGACAACGCTATGACTGTGAAAAACA 189
Db      66546  GSSWKAMTCTYAMMMMMWRMRAARWAAAGAGAGMGWRMRRTYWRYYRMRMRGRSRYYWA 66487
QY      190  TTGTATACGAAACCTAAAGCATCAACGAACTTGCTGTGTGGCGCTGAGCTACAC 249
Db      66486  MYAGYGRRRRGWYRCAYMMWMSRKATGKAMGAMYGWYTRWVKYKTKWTKKTKGYTCW 66427
QY      250  GTTGCAATATTGGTAAGCTATAAAAAATGCTTCGAGATAAGAAAAAGAAATTATTGAT 309
Db      66426  GYMMMKRMAATWMTTWTYYARWKYTKAATYKARWCMKCMWMMMKAKMSWMAWC 66367
QY      310  CCTTCACTATTTTCAAGCGGTCTGTAAACAGATACAGATGAGAAAAGCTTGCGTTAC 369
Db      66366  YGGYMAEMTWCASKSASKYRKGYSKYWKMSYSRGSWSTTYRARAARMMAARRRRWGR 66307
QY      370  ACCTCTGGCGATCTGCATCGTTCCTTTATGCTAATACTGACGAAATGTTATTAC 429
Db      66306  ATSRSKKGMKWKVAGMRMTMAARAMCAGSSWRMSASRWAGSKRRWSYTSYWCYAC 66247
QY      430  AGAGTCTCTAGTAATCTTACCTGTATGACACACATCTGATCAATACCAACTGTACTG 489
Db      66246  AMAAAMAMWMAAAATYRGYCKSMRGAMRGAMRWMAAMWRRAKCMRGGYKWCYRG 66187
QY      490  ACTAAA 495
Db      66186  ASRWWR 66181
```

Search completed: September 16, 2006, 15:58:07
Job time : 3723 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd

OM nucleic - nucleic search, using sw model

Run on: September 16, 2006, 14:41:19 ; Search time 514 Seconds
(without alignments)
7528.401 Million cell updates/sec

Title:	US-10-815-774-5
Perfect score:	555
Sequence:	1 ttggcgagtgtaatctctga.....atgtatccagctgtgcattt 555

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      5244920 segs, 3486124231 residues
Total number of hits satisfying chosen parameters: . 10489840
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

1:  N_Geneseq_8:*
2:  geneseqn1980s:*
3:  geneseqn1990s:*
4:  geneseqn2000s:*
5:  geneseqn2001as:*
6:  geneseqn2002as:*
7:  geneseqn2002bs:*
8:  geneseqn2002bs:*
9:  geneseqn2003as:*
10: geneseqn2003bs:*
11: geneseqn2003cs:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005as:*
15: geneseqn2005bs:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match.	Length	DB	ID	Description
1	555	100.0	555	4	AAE90280	AAE90280 Nucleotid
2	555	100.0	1080	4	AAE90281	AAE90281 Nucleotid
3	430.2	72.5	963	1	AE801515	AE801515 Corynef
4	344.2	62.0	555	3	AAZ49494	AAZ49494 Chryseoba
5	344.2	52.0	1080	3	AAZ49495	AAZ49495 Chryseoba
6	41.6	7.5	2000	11	ACU37108	ACU37108 Rice stre
7	39.8	7.2	11000	12	AD034435_3	AD034435_3 Cont
8	39	7.0	1105	12	AD022883	AD022883 Human so
9	39	6.7	2000	8	ADA71938	ADA71938 Rice gene
10	37	6.7	1585	4	AAK08825	AAK08825 Human dig
11	36.2	6.5	465	14	ACU58971	ACU58971 Human co
12	36.2	6.5	547	6	ABN61671	ABN61671 Human can
13	36	6.5	19576	6	ABN17057	ABN17057 Human can
14	36	6.5	19576	6	AAAG61258	AAAG61258 Human che
15	35.8	6.5	65359	6	AAAD6665	AAAD6665 Human tra
16	35.6	6.4	2483	3	AAZ90461	AAZ90461 E. canis
17	35.4	6.4	9300	11	ACU35887	ACU35887 Rice stre
18	35.4	6.4	9300	5	ABA15594	ABA15594 Human ner

19	35.2	6.3	2000	11	AC135363	AC135363 Rice stre
20	35	6.3	4283	4	ABH18719	ABH18719 Protophili
21	35	6.3	4590	5	AAH24065	AAH24065 Yeast AOD
C 22	35	6.3	106416	4	ABE118718	ABE118718 Protophili
C 23	34.8	6.3	1751	2	AAV40300	AAV40300 Sugarcbeet
24	34.8	6.3	2000	8	ADA71938	ADA71938 Rice gene
C 25	34.8	6.3	2000	11	AC137108	AC137108 Rice stre
26	34.8	6.3	50000	6	ABE56201	ABE56201 Ampyv gen
C 27	34.4	6.2	777	8	ACE73127	ACE73127 Steaphyloc
C 28	34.4	6.2	1072	2	AAV74746	AAV74746 Steaphyloc
C 29	34.4	6.2	1570	5	ABA18492	ABA18492 Human ner
C 30	34.4	6.2	3510	13	ADT47453	ADT47537 Bacterial
C 31	34.4	6.2	3943	4	AA831317	AA831317 Human CDN
C 32	34.4	6.2	3943	6	ABG66641	ABG66641 Human pol
C 33	34.4	6.2	3943	10	ADCI0663	ADCI0663 Human CDN
C 34	34.4	6.2	12582	4	AAK84357	AAK84357 Human imm
C 35	34.4	6.2	23821	3	AAA92469	AAA92469 Shewanell
C 36	34.2	6.2	2019	4	AAZ25737	AAZ25737 Tomato LI
37	34.2	6.2	4197	2	AAO99430	AAO99430 B. sphaer
C 38	34	6.1	511	13	ADQ57425	ADQ57425 Novel can
C 39	34	6.1	801	13	ADX64118	ADX64118 Plant fun
40	34	6.1	1653	13	ADTL16715	ADTL16715 Plant CDN
C 41	34	6.1	2286	10	ADB63717	ADB63717 Human CDH
C 42	34	6.1	2669	5	ABV24689	ABV24689 Human pro
C 43	34	6.1	2669	11	ACN91876	ACN91876 Breast ce
C 44	34	6.1	3212	10	ADC30112	ADC30112 Human not
C 45	34	6.1	3581	4	AAO9377	AAO9377 Human ves

ALIGNMENTS

	RESULT
ID	AAFP90280 standard; DNA, 555 BP.
AC	AAF90280;
D7	06-AUG-2003 (revised)
DT	22-AUG-2001 (first entry)
DE	Nucleotide sequence of a protein-deamidating enzyme.
KW	Protein-deamidating enzyme; mineral absorption; food allergy; dough; bakery; confectionary; ss.
OS	Chryseobacterium sp.
PN	EPI106696-Al.
PD	13-JUN-2001.
PF	04-DEC-2000; 2000EP-00310768.
PR	03-DEC-1999; 99JP-00345044.
PA	(AMAN-) AMANO ENZYME INC. Yamaguchi S;
DR	WP1; 2001-376907/40.
P-PDB:	AAB84386.
New enzyme	for use in e.g. bakery has an ability to deamidate amido groups in a protein.
Claim 8;	Page 22; 43pp; English.
The present	sequence encodes a protein-deamidating enzyme from Crysobacterium sp. number 9670. The enzyme is able to deamide amido groups in a protein by directly acting upon the amide groups without cutting peptide bonds and without cross-linking the protein. The enzyme thus reduces the mineral sensitivity of the protein and increases the

CC soluble mineral content in the protein-mineral solution, improving the
CC absorption of minerals in the human body. The enzyme reduces or removes
CC the toxicity and allergenic property of the protein in a food e.g. field
CC allergy. The enzyme is useful for the improvement of dough in the field
CC of bakery and confectionery e.g. for the production of crackers,
CC biscuits, cookies, pizza pies or crusts of pie; in producing soybean
CC protein products, in various food articles e.g. meat or fish products and
CC noodles; and for improving functionality of plant or animal protein.
CC (Updated on 06-AUG-2003 to correct OS field.)
XX

XX Sequence 555 BP; 166 A; 119 C; 107 G; 163 T; 0 U; 0 Other;

Query Match 100.0%; Score 555; DB 4; Length 555;

Best Local Similarity 100.0%; Pred. No. 1e-159; Mismatches 0; Indels 0; Gaps 0;

Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGAGTGAATTCCTGATGAGTACATTAATCTTATTCATCAATAAAGAT 60
DB 1 TTGGCGAGTGAATTCCTGATGAGTACATTAATCTTATTCATCAATAAAGAT 60

QY 61 CAGTCTTGCGGTAACCTTACGCGGCTCTCAACATGATCAATTCAGATTCCTGTAAC 120
DB 61 CAGTCTTGCGGTAACCTTACGCGGCTCTCAACATGATCAATTCAGATTCCTGTAAC 120

QY 121 GGATGTTATGCAAGAGCCCTAAGATGAGACAAATCTTATGAAACGCGCTATGACTGT 180
DB 121 GGATGTTATGCAAGAGCCCTAAGATGAGACAAATCTTATGAAACGCGCTATGACTGT 180

QY 181 GAAAAACAATTTGATACGAAACCTAAAGGATCAACAGAACTTGCTGTGAGCGTGG 240
DB 181 GAAAAACAATTTGATACGAAACCTAAAGGATCAACAGAACTTGCTGTGAGCGTGG 240

QY 241 AGCTACCAAGTGCATATTTGTAAGTATPAAAAATCTTCCGAGTAACGAAAAAGA 300
DB 241 AGCTACCAAGTGCATATTTGTAAGTATPAAAAATCTTCCGAGTAACGAAAAAGA 300

QY 301 ATTATGATCTCTTCACTATTTTCAAGCGGTCCTGTAACATACAGATGAGAAAGCT 360
DB 301 ATTATGATCTCTTCACTATTTTCAAGCGGTCCTGTAACATACAGATGAGAAAGCT 360

QY 361 TGGCTTAACAACCTTTCGAGATCGATCCGTTTCTCTTATGTAATACGACGAAT 420
DB 361 TGGCTTAACAACCTTTCGAGATCGATCCGTTTCTCTTATGTAATACGACGAAT 420

QY 421 GTTATTAACAAGAGTCTAGTAATCTTACCTGTATGACAAATCTGATCAATACCAAC 480
DB 421 GTTATTAACAAGAGTCTAGTAATCTTACCTGTATGACAAATCTGATCAATACCAAC 480

QY 481 TGCTACTGACTAATTTTCACTGCTTCCGAGATGTTCTCTTCAACCGACCGGATGTA 540
DB 481 TGCTACTGACTAATTTTCACTGCTTCCGAGATGTTCTCTTCAACCGACCGGATGTA 540

QY 541 TCCAGCTGTGATTT 555
DB 541 TCCAGCTGTGATTT 555

RESULT 2
AAF90281 ID AAF90281 standard; DNA; 1080 BP.

XX AAF90281;
XX 06-AUG-2003 (revised)
XX 22-AUG-2001 (first entry)

DE Nucleotide sequence of a protein-deamidating enzyme.

XX Protein-deamidating enzyme; mineral absorption; food allergy; dough;
XX bakery; confectionery; ss.

OS Chryseobacterium sp.

FH Key Location/Qualifiers

FT CDS 61..1023

FT /tag= a /product= "protein-deamidating enzyme"

XX EP110696-A1.

PD 13-JUN-2001.

XX 04-DEC-2000; 2000EP-00310768.

XX 03-DEC-1999; 99UP-00345044.

XX (AMANO-) AMANO ENZYME INC.

XX Yamaguchi S;

XX WPI; 2001-376907/40.

XX P-PADB; AAB84387.

XX New enzyme for use in e.g. bakery has an ability to deaminate amido groups in a protein.

XX Example 11; Page 23; 43p; English.

CC The present sequence encodes a protein-deamidating enzyme from
CC Chryseobacterium sp. number 9670. The enzyme is able to deaminate amido
CC groups in a protein by directly acting upon the amido groups without
CC cutting peptide bonds and without cross-linking the protein. The enzyme
CC thus reduces the mineral sensitivity of the protein and increases the
CC soluble mineral content in the protein-mineral solution, improving the
CC absorption of minerals in the human body. The enzyme reduces or removes
CC the toxicity and allergenic property of the protein in a food e.g.
CC allergy. The enzyme is useful for the improvement of dough in the field
CC of bakery and confectionery e.g. for the production of crackers,
CC biscuits, cookies, pizza pies or crusts of pie; in producing soybean
CC protein products, in various food articles e.g. meat or fish products and
CC noodles; and for improving functionality of plant or animal protein.
CC (Updated on 06-AUG-2003 to correct OS field.)
XX

XX Sequence 1080 BP; 371 A; 204 C; 203 G; 302 T; 0 U; 0 Other;

Query Match 100.0%; Score 555; DB 4; Length 1080;

Best Local Similarity 100.0%; Pred. No. 1.4e-159; Mismatches 0; Indels 0; Gaps 0;

Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGAGTGAATTCCTGATGAGTACATTAATCTTATTCATCAATAAAGAT 60
DB 466 TTGGCGAGTGAATTCCTGATGAGTACATTAATCTTATTCATCAATAAAGAT 525

QY 61 CAGTCTTGCGGTAACCTTACGCGGCTCTCAACATGATCAATTCAGATTCCTGTAAC 120
DB 526 CAGTCTTGCGGTAACCTTACGCGGCTCTCAACATGATCAATTCAGATTCCTGTAAC 120

QY 121 GGATGTTATGCAAGAGCCCTAAGATGAGACAAATCTTATGAAACGCGCTATGACTGT 180
DB 121 GGATGTTATGCAAGAGCCCTAAGATGAGACAAATCTTATGAAACGCGCTATGACTGT 180

QY 181 GAAAAACAATTTGATACGAAACCTAAAGGATCAACAGAACTTGCTGTGAGCGTGG 240
DB 586 GAAATGATTTGATACGAAACCTAAAGGATCAACAGAACTTGCTGTGAGCGTGG 240

QY 241 AGCTACCAAGTGCATATTTGTAAGTATPAAAAATCTTCCGAGTAACGAAAAAGA 300
DB 706 AGCTACCAAGTGCATATTTGTAAGTATPAAAAATCTTCCGAGTAACGAAAAAGA 300

QY 301 ATTATGATCTCTTCACTATTTTCAAGCGGCTCTGTAACAGATCAGCATGAGAAAGCT 360
DB 766 ATTATGATCTCTTCACTATTTTCAAGCGGCTCTGTAACAGATCAGCATGAGAAAGCT 360

QY 361 TGGCTTAACAACCTTTCGAGATCGATCCGTTTCTCTTATGTAATACGACGAAT 420
DB 826 TGGCTTAACAACCTTTCGAGATCGATCCGTTTCTCTTATGTAATACGACGAAT 420

QY 421 GTTATTACGAGAGCTCTAGTAATCTTACCTGTATGACAAATCTGATCAATACCAC 480
 DB 886 GTTATTACGAGAGCTCTAGTAATCTTACCTGTATGACAAATCTGATCAATACCAC 945
 QY 481 TGTGTACTGACTAATTTTACAGCTCTTCCGGAGTGTCTTCCTTACCTGACCGGATGTA 540
 DB 946 TGTGTACTGACTAATTTTACAGCTCTTCCGGAGTGTCTTCCTTACCTGACCGGATGTA 1005
 QY 541 TCCAGCTGTGATTT 555
 DB 1006 TCCAGCTGTGATTT 1020

RESULT 3
 ABE01515
 ID ABE01515 standard; DNA; 963 BP.
 XX ABE01515;
 AC ABE01515;
 XX 26-JAN-2006 (first entry)
 DT
 XX Coryneform protein production system-related DNA, SEQ ID NO:3.
 DE
 XX protein production; cell culture; genetically engineered microorganism;
 KW ds; gene.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT 1..963
 FT CDS /tag= a
 FT /product= "Undefined protein"
 PN MO2005103278-A1.
 XX
 PD 03-NOV-2005.
 PF 20-APR-2005; 2005MO-JP007518.
 XX
 -PR 20-APR-2004; 2004JP-00124196.
 PR 13-JUN-2005; 2005JP-00005896.
 XX
 PA (AJIN) AJINOMOTO CO INC.
 PI Date M, Kikuchi Y, Itaya H, Nakamura N;
 XX WPI; 2005-748042/76.
 DR P-PSDB; ABE01516.
 XX
 PT Producing heterologous protein by culturing Coryneform bacterium having
 PT expression gene construct comprising promoter sequence, nucleic acid
 PT encoding Tat system-dependent signal peptide region and nucleic acid
 PT encoding foreign protein.
 PS
 PS Disclosure; SEQ ID NO 3; 47pp; Japanese.
 XX
 CC The new invention relates to a method of producing heterologous protein,
 CC by culturing a Coryneform bacterium having an expression gene construct
 CC carrying, in the direction from the 5'-terminal to 3'-terminal, a
 CC promoter sequence functioning in the bacterium, a nucleic acid sequence
 CC encoding a Tat system-dependent signal peptide region and a nucleic acid
 CC sequence encoding a foreign protein, and allowing the bacterium to
 CC produce and secrete the protein. The signal peptide of (M1) comprises SEQ
 CC ID NO: 28-32. The signal peptide is a signal peptide of isomaltose dectran,
 CC which comprises SEQ ID NO: 6, or of trimethylamine N-oxidoreductase,
 CC which comprises SEQ ID NO: 8. The bacteria comprises gene that encodes
 CC Tat system secretion peptide such as tatA, tatB, tatC or tatZ. (M1) is
 CC useful for producing heterologous protein e.g. glutaminase or
 CC isomaltodextran. (M1) enables efficient production of heterologous
 CC protein. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences. The present

CC sequence is a DNA related to the invention.
 XX
 SX Sequence 963 BP; 240 A; 290 C; 207 G; 226 T; 0 U; 0 Other;
 QY
 Query Match 77.5%; Score 430.2; DB 14; Length 963;
 Best Local Similarity 85.9%; Pred. No. 2.8e-121;
 Matches 477; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 1 TTGGGAGTGTAAATTCCTGATGTAGTACATTAATTTCTTTATTCATCAATTAAGAAT 60
 DB 406 TTGGGCTCCGTCATTCAGATGTGGTACACCTCCCACTCTCTTACCAAAATCAAGAAC 465
 QY 61 CAGTCTTGGGTAACCTTACAGCGCTCTCAACCATGATCAATTCAGATATCTGTAGAC 120
 DB 466 CAGTCTTGGGTAACCTTACAGCGCTCTCAACCATGATCAATTCAGATATCTGTAGAC 525
 QY 121 GGATGTTATGCAAGAGCCCATTAAGATGAGCAAAATTTAATGAACAACGGCTATAGCT 180
 DB 526 GGCTGCTACGACGGGCCCAAGATGGCCAGATCTGATGAACAACGGCTATAGCTAT 585
 QY 181 GAAAAACAATTTGTATACGAAACCTTAAGGATCAACAGAACTGTGTGGCGTGG 240
 DB 586 GAAAGCAATTCGTGTACGATTAACCTCAAGGATCAACGGCACTGCTGCTGGCGTGG 645
 QY 241 AGCTACCAAGTTCGAATATTTGTAAGCTATTAATAATGCTTCGGAGTAAACGAAAAAGA 300
 DB 646 AGCTACCAAGTTCGAATATTTGTAAGCTATTAATAATGCTTCGGAGTAAACGAAAAAGC 705
 QY 301 ATTATTTGATCTCTTACATATTTTCAAGCGGTCTCTGTAACAGATACAGATGAAGAACGT 360
 DB 706 ATTATTTGATCTCTTACATATTTTCAAGCGGTCTCTGTAACAGATACAGATGAAGAACGT 765
 QY 361 TCGGTTAACACCTCTTGGGATCTGATCCGTTTCTCTTATGCTAATATCAGAGAAAT 420
 DB 766 TCGGTTAACACCTCTTGGGATCTGATCCGTTTCTCTTATGCTAATATCAGAGAAAT 825
 QY 421 GTTATTACGAGAGCTCTAGTAATCTTACCTGTATGACAAATCTGATCAATACCAC 480
 DB 826 GTTATTACGAGAGCTCTAGTAATCTTACCTGTATGACAAATCTGATCAATACCAC 885
 QY 481 TGTGTACTGACTAATTTTACAGCTCTTCCGGAGTGTCTTCCTTACCTGACCGGATGTA 540
 DB 886 TGTGTACTGACTAATTTTACAGCTCTTCCGGAGTGTCTTCCTTACCTGACCGGATGTA 945
 QY 541 TCCAGCTGTGATTT 555
 DB 946 TCCAGCTGTGATTT 960

RESULT 4
 AA249494
 ID AA249494 standard; DNA; 555 BP.
 XX AA249494;
 AC AA249494;
 XX
 DT 15-SEP-2003 (revised)
 DT 04-APR-2000 (first entry)
 XX
 DE Chryseobacterium gleum protein-deamidating enzyme encoding DNA.
 KW Protein deamidating enzyme; soil bacterium; deamidation activity;
 KW amido group; carboxyl group; ammonia; cross linking; peptide bond;
 KW protein engineering; surface hydrophobicity; toxicity; allergic;
 KW mineral sensitivity; calcium; absorption; mineral enhancing agent; ds.
 XX Chryseobacterium gleum; 'JCM 2410'.
 OS
 XX
 FH Key Location/Qualifiers
 FT 1..555
 FT /tag= a
 FT /product= "Protein deamidating enzyme"
 FT /function= "Deamidate amido groups in a protein"
 FT /note= "Improves protein function"

```

XX  EP976829-A2.
XX  PD 02-FEB-2000.
XX  PF 04-JUN-1999; 99EP-00304367.
XX  PR 04-JUN-1998; 98JP-00173940.
XX  (AMANO ) AMANO PHARM KK.
XX  Yamaguchi S, Matsuura A;
XX  WPI; 2000-118552/11.
XX  P-PSDB; AAY44582.
XX  New enzyme for modifying and improving the function of proteins and/or
XX  peptides has deamidating activity without causing cross linking.
XX  Claim 8; Page 23; 57pp; English.
XX  The present sequence is the DNA encoding the protein-deamidating enzyme,
XX  isolated from a new strain of soil bacterium, Chryseobacterium gleum JCM
XX  2410. The enzyme exerts the deamidation activity by directly acting upon
XX  side chain amido groups in the protein in bonded state and releasing side
XX  chain carboxyl groups and ammonia. It can deaminate high molecular weight
XX  proteins, without cross linking and cleavage of peptide bonds, to improve
XX  protein function. This sequence is used for protein engineering, to cause
XX  an increase in surface hydrophobicity and improve the function of a plant
XX  or animal protein. It can also be used to remove or reduce toxicity of
XX  allergenic property of proteins in food, decrease mineral sensitivity of
XX  protein, to allow greater absorption into the body and to solubilise
XX  calcium for use in drinks and mineral enhancing agents. (Updated on 15-
XX  SBP-2003 to standardise OS field)
XX  Sequence 555 BP; 159 A; 122 C; 112 G; 162 T; 0 U; 0 Other;
XX  Query Match 62.0%; Score 344.2; DB 3; Length 555;
XX  Best Local Similarity 76.7%; Pred. No. 5.6e-95;
XX  Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
QY 7 AGTGAATTCCTCGAATGTAGCTACATTAATTTCTTATTCATCAATAAATACAGTCT 66
DB 7 AGGTGTAATTCCTGATCTGCAACGCTGACAGTTTATTTACCGATGAAAACCAAGCT 66
QY 67 TCGCGTACCTCTACGCGCTCTCTCAACATGATCAATTCAGATATCTCTGACGAGATGT 126
DB 67 TCGGGAATCTTACAGCATCTTCTCTGTATCACTTCAAGTATCCGGTTGACGAGATGT 126
QY 127 TATGCAAGAGCCCATAGATGAGACAAATCTTAATGAAACAGGCTATGATCTGGAAGAA 186
DB 127 TATGCAAGGGCTCACAAAATAGACAAATCTTATTAAGACGCGCTATGATCTGGAAGAA 186
QY 187 CAATTTGATACGGAACCTTAAGGCATCAACAGGAATCTTGCTGTGCTGAGACTAC 246
DB 187 CAGTTCGATATGATGATGAGAGCTTCTACAGAAATGCTGTGATCATGGTATAT 246
QY 247 CACGTTGCAATATGTGTAAGCTATATAAAATGCTTCGAGATACGGAAGAAATTAAT 306
DB 247 CACGTAGCAATTTTGTGTAAGCTTCAAAAATGCTTCAGGAATTTGAAAAAAGATCAATA 306
QY 307 GATCTCTCACTATTTTCAAGCGGCTCTGTAACAGATACAGCTGAGAAACGCTTGCGTT 366
DB 307 GATCTCTCACTATTTTCCACGCGTCTCTGTAACAGATTCGAGAGAGCTGATGATACC 366
QY 367 AACACCTCTTGCGGATCTGCATCCGTTCTCTTAATGTAATATGCGAGAAATGTTAT 426
DB 367 AACACAGCTGCGGATCTGCTGTGATCTTCTACAGCCAAATACAGCAAGAAATGTTTAC 426
QY 427 TACAGAAAGTCTAGTAATTTCTTAACCTGTATGACAAACATCTGATCAATACCACTGTGTA 486
DB 427 TACAGAAAGTCCGTCAGGTCATTAATCTGTATGATTAACAACTAATGTGAATCAATTTGTGTA 486

```

```

QY 487 CTGACTAATTTTCACTGCTTCCGAGATGTTCTCTTCACCTGACCGAGATGATCCAGC 546
DB 487 TTAACATATTTCTCATCCCTTTACAGATGTTCTCTTCCACGACCAAGATGACAGC 546
QY 547 TGTGGAATTT 555
DB 547 TGTGGAATTT 555
RESULT 5
AAZ49495
AAZ49495 standard; DNA; 1080 BP.
AAZ49495;
AC 15-SEP-2003 (revised)
DT 04-APR-2000 (first entry)
XX Chryseobacterium gleum protein-deamidating prepro-enzyme encoding gene.
XX KM protein deamidating prepro-enzyme; soil bacterium; deamidation activity;
XX KM amido group; carboxyl group; ammonia; cross linking; peptide bond;
XX KM protein engineering; surface hydrophobicity; toxicity; allergic;
XX KM mineral sensitivity; calcium; absorption; mineral enhancing agent; ds.
XX OS Chryseobacterium gleum; 'JCM 2410'.
XX FH Key Location/Qualifiers
XX FH CDS 61..1020
XX FT /*tag= a
XX FT /product= "C. gleum protein deamidating prepro-enzyme"
XX FT /function= "deaminate amido groups in a protein"
XX FT 61..123
XX FT /*tag= b
XX FT /note= "Corresponds to pre region of the protein"
XX FT mat_peptide 463..1017
XX FT /*tag= c
XX FT /label= "Mature protein deamidating enzyme"
XX FT /note= "Improves protein function"
XX PN EP976829-A2.
XX PD 02-FEB-2000.
XX PF 04-JUN-1999; 99EP-00304367.
XX PR 04-JUN-1998; 98JP-00173940.
XX PA (AMANO ) AMANO PHARM KK.
XX PI Yamaguchi S, Matsuura A;
XX DR WPI; 2000-118552/11.
XX DR P-PSDB; AAY44583.
XX PT New enzyme for modifying and improving the function of proteins and/or
XX PT peptides has deamidating activity without causing cross linking.
XX PS Example 26; Page 25; 57pp; English.
XX CC The present sequence is the gene encoding the protein-deamidating prepro-
XX CC enzyme, isolated from a new strain of soil bacterium, Chryseobacterium
XX CC gleum JCM 2410. The enzyme exerts the deamidation activity by directly
XX CC acting upon side chain amido groups in the protein in bonded state and
XX CC releasing side chain carboxyl groups and ammonia. It can deaminate high
XX CC molecular weight proteins, without cross linking and cleavage of peptide
XX CC bonds, to improve protein function. This sequence is used for protein
XX CC engineering, to cause an increase in surface hydrophobicity and improve
XX CC the function of a plant or animal protein. It can also be used to remove
XX CC or reduce toxicity of allergenic property of proteins in food, decrease
XX CC mineral sensitivity of protein, to allow greater absorption into the body
XX CC and to solubilise calcium for use in drinks and mineral enhancing agents.
XX CC (Updated on 15-SBP-2003 to standardise OS field)

```

XX SQ Sequence 1080 BP, 354 A, 210 C, 205 G, 311 T, 0 U, 0 Other;
 Query Match 62.0%; Score 344.2; DB 3; Length 1080;
 Best Local Similarity 76.7%; Pred. No. 7.6e-95;
 Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 7 AGTGAATTCCTGATAGTACATTAATTTCTTATTCATCAATTAAGATCAGTCT 66
 DB 469 AGTGAATTCCTGATAGTACATTAATTTCTTATTCATCAATTAAGATCAGTCT 528
 QY 67 TGGGTAACCTTACCGGCTCTCAATGATGATCAATTCATGATGATGATGATGAT 126
 DB 529 TGGGTAACCTTACCGGCTCTCAATGATGATGATGATGATGATGATGATGAT 588
 QY 127 TATGCAAGAGCCCATAGATGAGACAAATCTTAATGAAACGGCTATGCTGAAAA 186
 DB 589 TATGCAAGAGCCCATAGATGAGACAAATCTTAATGAAACGGCTATGCTGAAAA 648
 QY 187 CAATTTGTATACGAAACCTTAAGGATCAACAGAACTGCTGTGCTGAGCTAC 246
 DB 649 CAATTTGTATACGAAACCTTAAGGATCAACAGAACTGCTGTGCTGAGCTAC 708
 QY 247 CACGTTGCAATTTGTTAGTAAATGCTTCCGAGTAAACGAAAAAGATTTAT 306
 DB 709 CACGTTGCAATTTGTTAGTAAATGCTTCCGAGTAAACGAAAAAGATTTAT 768
 QY 307 GATCTCTCAATTTTCAAGCGGCTCTGTAACAGATGAGCAAGGATGCTGCTT 366
 DB 769 GATCTCTCAATTTTCAAGCGGCTCTGTAACAGATGAGCAAGGATGCTGCTT 828
 QY 367 AACACCTTGGCGATCTGATCCGTTCTCTTATGCTAATCTCAGAAATGTTAT 426
 DB 829 AACACCTTGGCGATCTGATCCGTTCTCTTATGCTAATCTCAGAAATGTTAT 888
 QY 427 TACAGAGTCTGATGATTTCTTACCTGATGACAAACATCTGATGATGAT 486
 DB 889 TACAGAGTCTGATGATTTCTTACCTGATGACAAACATCTGATGATGAT 948
 QY 487 CTGACATAATTTTCACTGCTTCCGATGCTTCTCTTCACTGACGAGATGCTCAG 546
 DB 949 CTGACATAATTTTCACTGCTTCCGATGCTTCTCTTCACTGACGAGATGCTCAG 1008
 QY 547 TGTGATTT 555
 DB 1009 TGTGATTT 1017

RESULT 6
 ID ACL37108 standard; cDNA; 2000 BP.
 XX ACL37108;
 AC 02-JUN-2005 (first entry)
 XX
 DE Rice stress-regulated promoter SEQ ID NO:15671.
 XX
 KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
 XX agriculture.
 XX
 OS Oryza sativa.
 XX
 PN MO2003008540-A2.
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019668.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 XX
 PR 24-AUG-2001; 2001US-0314662P.
 XX
 PR 26-SEP-2001; 2001US-0325277P.
 XX
 PR 21-NOV-2001; 2001US-0332132P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA
 XX
 PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
 PI Moughamer T, Provart N, Riecke D, Zhu T,
 XX WPI; 2003-248011/24.
 DR
 XX
 PT New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g., cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 XX
 PS Claim 48; SEQ ID NO 15671; 89pp; English.
 XX
 CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention

XX SQ Sequence 2000 BP, 510 A, 350 C, 268 G, 518 T, 0 U, 354 Other;
 Query Match 7.5%; Score 41.6; DB 11; Length 2000;
 Best Local Similarity 10.4%; Pred. No. 0.084;
 Matches 43; Conservative 194; Mismatches 173; Indels 4; Gaps 1;

QY 111 TCCTGTAGACGATGTTATGCAAGCCCATAGATGAGCAATCTTATGACACCG 170
 DB 40 KYKMYYYRTCSYKRWMTSMKRYAYMMRWKMMWMMVYMGMYVGYMMWMTWYSR 99
 QY 171 CTATGACGTGTGAAAAACAATTTGTATACGAAACCTTAAGCATGACAGAACTTCTG 230
 DB 100 TMTYTAAYKMKMMRRKMSRMAAMWYSWYRTSRMTGTRKKRMMRMYMTWYWC 159
 QY 231 TGTGCGTGGAGCTACACGTTGCAATATGTTAGTAAATGCTCCGAGTAAC 290
 DB 160 YTMSSMR---YRWRKMKRAYKMAATTTTMRKCMWMTCTGYKMAAMASAR 215
 QY 291 GGAATAAGAAATTAATGATCTTCACTATTTTCAAGCGCTCTGTAACAGATACAGCAT 350
 DB 216 ARRAMMMRRKMSMTWATWMSWMSWMTWKRMGMWMAVSSYTMWMMWMTCTRYAWY 275
 QY 351 GAGAAACGCTCCGTTAACACCTCTTGGGATGCTGATCCGTTCTCTTAATGATAC 410
 DB 276 KKKAAATMTRCTWYMMMAAMTSAMRAMCAGWASVATKYAMWTTMKRRRRRYACMAWR 335
 QY 411 TGCAGAAATGTTTATACAGAGTCTAGTATCTTACCTGATGACCAACATCTGAT 470
 DB 336 WTKMSMRKTKMWTMAARSWCKKMMMAKAMATMCAATYSTRYAAARRRWRMSWA 395
 QY 471 CAATACCACTGTGATGACTAAATTTTCACTGCTTCCGAGATTTCTCTCT 524
 DB 396 AWYSAKYVYAKSYWYMKARWMTTYVAYAAAAATRTAAVAGRTWMTWCTM 449

RESULT 7
 ADO34435_3
 Continuation (4 of 7) of ADO34435 from base 300001 (Human SLIT-3 genomic sequence.)
 WP Sequence split into 7 fragments LOCUS ADO34435 Accession ADO34435
 WP Fragment Name Begin End
 WP ADO34435_0 1 110000
 WP ADO34435_1 100001 210000
 WP ADO34435_2 200001 310000
 WP ADO34435_3 300001 410000

WP AD034435_4 400001 510000
 XX AD034435_5 500001 610000
 WP AD034435_6 600001 634886

Query Match 7.2%; Score 39.8; DB 12; Length 110000;
 Best Local Similarity 49.3%; Pred. No. 1.7;
 Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 18 TGATGAGCTACATTAATCTTATTCATCAATAAAGATCAGTCTTGCGGTACCTC 77
 DB 85206 TGGTTGTTTAACTAATCTATACACAGTTTCATATGACATTTACTTGTAGGCTCC 85265
 QY 78 TACGGCGCTCTCAACATGATCATCATTCAGATATCTCTGAGACGATGTTATCCAGAGC 137
 DB 85266 TTTTGTGCCAAAGCATGCGGTTCCATACAGCAGCCTCTGACACATGTAGCTGATGAGC 85325
 QY 138 CCATAGATGAGACAAATCTTAAATGACCAACGGTATGACTGTGAAAAACAATTTGATA 197
 DB 85326 AGATTAATATATGGCTAGTCCAAATTAAGTATGCTGTAAAGTATTAATGCAAAAGTGATT 85385
 QY 198 CGGAACCTAAAGGATCAACAGGAACTTGC 228
 DB 85386 TCCAAACAGATTTCAAAAAAGAAATGC 85416

RESULT 8
 ADQ22883
 ID ADQ22883 standard; DNA; 1105 BP.
 AC ADQ22883;
 XX 26-AUG-2004 (first entry)
 DT
 XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 5703.
 DE
 XX soft tissue sarcoma; cyrostatic; gene therapy; vaccine; screening; human;
 KM ds.
 XX Homo sapiens.
 OS
 XX MO2004048938-A2.
 PN
 XX 10-JUN-2004.
 PD
 XX 26-NOV-2003; 2003WO-US038193.
 PF
 XX 26-NOV-2002; 2002US-0429739P.
 PR
 XX (PROT-) PROTEIN DESIGN LABS INC.
 XX
 PA Aziz N, Ginsburg WM, Zlotnik A;
 PI WPI; 2004-441208/41.
 DR
 XX Early detection of soft tissue sarcoma comprises determining expression
 PT of a gene in a first soft tissue sample and a normal soft tissue sample
 PT and comparing the gene expression, also useful in treating soft tissue
 PT sarcoma.
 XX
 PS Example 2; SEQ ID NO 5703; 210bp; English.
 XX
 CC The invention relates to a novel method for detecting soft tissue sarcoma
 CC which comprises obtaining a first soft tissue sample from an individual,
 CC and a normal soft tissue sample from the same or different individual,
 CC determining the expression of a gene in both samples and comparing the
 CC expression of the gene in both soft tissue samples, where a higher level
 CC of protein expression in the first soft tissue sample indicates the
 CC presence of soft tissue sarcoma. The method of the invention has
 CC cyrostatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC DNA of the invention. The current sequence is not shown within the

CC specification per se but was submitted in CD format by the inventor.
 XX Sequence 1105 BP; 355 A; 204 C; 212 G; 334 T; 0 U; 0 Other;
 SQ

Query Match 7.0%; Score 39; DB 12; Length 1105;
 Best Local Similarity 52.8%; Pred. No. 0.41;
 Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 182 AAAAACAATTTGTATACGGAACCTAAAGCATCAACAGAACTTGCTGTGCGGTGCA 241
 DB 91 AAAAGAGTTTATAGAAATCTGGTTACTTCTTGAATATAGAACTTATAGATTTAA 150
 QY 242 GCTACACAGTTGCAATATTTGTAAGCTATTAATAATGCTTCGAGTAAAGGAAAAA 301
 DB 151 GCTACTGTTTATATATGAGGTTAGGCTVAGAGTACCMAAGGAAAAAGGTTAAATTG 210
 QY 302 TTATGATCTTCACTATTTTCAACGGTCTCTGTAACG 340
 DB 211 CTAAAGAAATTTCAACMAATTTAAGTTGTCTTTAAAAG 249

RESULT 9
 ADA71938/C
 ID ADA71938 standard; DNA; 2000 BP.
 AC ADA71938;
 XX 20-NOV-2003 (first entry)
 DT
 XX Rice gene, SEQ ID 5263.
 DE
 XX Plant; bacterial infection; fungal infection; viral infection; rice;
 KM gene; ds.
 XX Oryza sativa.
 OS
 XX WO2003000898-A1.
 PN
 XX 03-JAN-2003.
 PD
 XX 22-JUN-2001; 2001WO-IB001105.
 PF
 XX 22-JUN-2001; 2001WO-IB001105.
 PR
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.
 DR
 XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 PS Claim 27; SEQ ID NO 5263; 899bp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
 XX
 CC Query Match 6.7%; Score 37; DB 8; Length 2000;
 CC Best Local Similarity 8.2%; Pred. No. 2.2;

PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250319P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-502630/55.
 XX
 PT polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognosing disorders of the
 PT digestive system, particularly cancer and cancer metastases.
 XX
 PS Disclosure; SEQ ID NO 4401; 966bp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a genomic DNA fragment
 CC encoding a digestive system antigen of the invention
 XX
 SQ Sequence 15865 BP; 4775 A; 3204 C; 3309 G; 4577 T; 0 U; 0 Other;

Query Match 6.7%; Score 37; DB 4; Length 15865;
 Best Local Similarity 50.9%; Pred No. 5.3;
 Matches 88; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
 QY 54 AAAGATCAGCTTGGTGGTACCTTACGCGCTCCACCATCATCATTCAGATATCC 113
 DB 13604 NATGACTAGCTTCAAGGATCTCTTGGTACTGAAAACATTCATTCGAAAAA 13545
 QY 114 TGTAGACGATTTATGCAAGGCCCTTAAGTGAACAAATCTTAATCAACAGCGCTA 173
 DB 13544 AACAAAAAACTTTCATGGAAGAAATTAATTAGAGAAATCATTAATGCTTAACGCTG 13485
 QY 174 TGACTGTGAAAAACATTTGTATACGGAACCTTAAGGCATCAACAGAACTT 226
 DB 13484 GTACAGAGGAAAAATTTGTAAATTAAGATATAGTAATGTACAGACTT 13432
 RESULT 11
 ID ACL58971 standard; cDNA; 464 BP.
 AC ACL58971;
 XX 24-MAR-2005 (first entry)
 DE Human colon cancer differentially expressed polynucleotide, SEQ ID:5106.
 XX
 KW Differential expression; diagnosis; therapy; drug screening; cancer;
 KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
 KW ss.
 XX Homo sapiens.
 OS
 PN WO200500087-A2.
 PD 06-JAN-2005.
 PF 13-MAY-2004; 2004WO-US015421.
 XX
 PR 03-JUN-2003; 2003US-0475872P.
 PA (CHIR) CHIRON CORP.
 PI Randazzo F, Moler E, Escobedo J, Garcia PD;
 XX WPI; 2005-075421/08.
 DR
 XX
 PT New isolated polynucleotides, which are differentially expressed in colon
 PT cancer cell, useful for treating cancer, e.g. colon cancer, breast
 PT cancer, or pancreatic cancer.
 XX
 PS Claim 1; SEQ ID NO 5106; 97pp; English.
 XX
 CC The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
 CC are differentially expressed in colon cancer cells. The invention also
 CC relates to vectors and host cells comprising a differentially expressed
 CC polynucleotide of the invention; a method for detecting a cancerous cell
 CC by detection of a gene product of the polynucleotides; a method for
 CC inhibiting a cancerous phenotype of a cell by inhibiting a gene product
 CC of the polynucleotides; a method of treating an individual with cancer by
 CC administration of a modulator of a gene product of the polynucleotides;
 CC and an isolated antibody that specifically binds to a polypeptide encoded
 CC by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
 CC antibodies, and methods are useful for the detection of cancerous cells;
 CC for the diagnosis, prognosis and management of cancer; for the
 CC identification of agents that modulate the phenotype of cancerous cells;
 CC for the identification of therapeutic targets for cancer chemotherapy;
 CC and for the treatment of cancer, especially colon cancer and metastasized
 CC colon cancer, but also breast or pancreatic cancer. The polynucleotides
 CC are also useful as a source of probes or primers for use in diagnostic
 CC methods. The differentially expressed polynucleotides or their encoded
 CC proteins can additionally be used as vaccines to modulate primary immune

Db	6024	TTCTGTAATTTAAATTAATCACTAATTAATTAATAAACAATCCAAATATTCACCT	5965
Qy	372	CTCTGGCGAGTCGATCCGTTTCCCTGTATGCTAATATACGAGAAATGTTATTAACG	431
Db	5964	AACTAAAAACATTTAACTTTTTCGCTAAAAATATATATATATATTTATTAATAAA	5905
Qy	432	AAAGTCTAGTAATTTCTTACCTGTATGACAAACAATCTGATCAATACCACTGTACTGAC	491
Db	5904	AAAAATAACTCCTTTTAAACCAAAATATTAATATTTTCTAAATATTAATCTTACTATCCAC	5845
Qy	492	TAAATTTT 499	
Db	5844	AAAAATTT 5837	
RESULT 14			
AAS61258/c			
ID	AAS61258	standard; DNA; 19576 BP.	
XX	AAS61258;		
XX	29-JAN-2002	(first entry)	
XX			
DE		Human gene regulation-associated gene oligonucleotide #213.	
XX			
KW		Human; Gene regulation-associated gene; severe combined immunodeficiency;	
KW		cardiac damage; inflammatory response; Haemophilia; Werner syndrome;	
KW		asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;	
*KW		renal disease; Precocious; cardiac allograft vascular disease;	
KW		colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;	
KW		immunostimulant; caridian; antiinflammatory; coagulant; antiasthmatic;	
KW		nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.	
XX			
XX	Homo sapiens.		
XX			
PN	WO200177375-A2.		
PD	18-OCT-2001.		
XX			
PF	06-APR-2001; 2001WO-EP003968.		
XX			
PR	06-APR-2000; 2000DE-01019058.		
PR	07-APR-2000; 2000DE-01019173.		
PR	30-JUN-2000; 2000DE-01032529.		
PR	01-SEP-2000; 2000DE-01043826.		
XX			
PA	(EPIC-) EPIGENOMICS AG.		
XX			
PI	Olek A, Piepenbrock C, Berlin K;		
XX			
DR	WPI; 2002-017470/02.		
XX			
PT	New nucleic acid sequences from chemically modified genes associated with		
PT	gene regulation, useful for analyzing cytosine methylations for diagnosis		
PT	and therapy of diseases e.g. severe combined immunodeficiency disease.		
XX			
PS	Claim 1; SEQ ID NO 219; 26bp; English.		
XX			
CC	The invention relates to 224 nucleic acid sequences comprising at least		
CC	18 bases of a chemically pretreated gene associated with gene regulation		
CC	selected from 43 known genes (or complementary sequences). The chemical		
CC	pretreatment converts cytosine bases unmethylated at the 5-position to		
CC	uracil or another base with hybridisation behaviour dissimilar to		
CC	cytosine, to enable analysis of cytosine methylations. The DNA sequences,		
CC	(or sets/arrays) and method are useful in the diagnosis of		
CC	diseases (or predisposition to diseases) associated with gene regulation		
CC	and in therapy of such diseases, by enabling analysis of the cytosine		
CC	methylation patterns of such genes, kits are provided. They are		
CC	especially useful in diagnosis and therapy of e.g. severe combined		
CC	immunodeficiency disease, cardiac disorders, haemophilia, solid tumours		
CC	and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen		
CC	syndrome, renal disease, precocious, graft versus-host disease. The		
CC	present sequence is a sequence included in the sequence data for this		

CC	specification and is associated with the human gene regulation-associated
CC	genes. Note: The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
5Q	Sequence 19576 BP; 5357 A; 307 C; 4161 G; 9750 T; 0 U; 1 Other;
	Query Match 6.5%; Score 36; DB 6; Length 19576;
	Best Local Similarity 49.5%; Pred. No. 12;
	Matches 93; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY	
312	TTCACTATTTTCACGCGGTCTGTATACAGATACGATGACGAGAAACCTTGCCTTAACAC 371
6024	TTCTATATTTTAAATTAATCTACTACCTATTAATAATTAACATCCATAATTCTACCT 5965
DB	
372	CTCTGGGAGATCTGATCCGTTTCCCTTATGCTAATATCTGCAGAAATGTTATTAACAG 431
5964	AACTAATAAACATTTAACTTTTGGCTAAATAATATATATATATATATATATATATAAAA 5905
QY	
432	AAGTCTGATGTAATTTCTTACTCTGTATGACACAAATCTGATCAATACCACTGTACTGCAC 491
5904	AAAAATATCTCTTTTATACCAAAATATTAATATTTCTAATATTAATATCTACTATCCAC 5845
DB	
492	TAAATTTT 499
5844	AAAAATTT 5837
RESULT 15	
AAD46665/C	
ID	AAD46665 standard; DNA; 65359 BP.
XX	
AC	AAD46665;
XX	
DT	27-JAN-2003 (first entry)
XX	
DE	Human transporter gene.
XX	
KW	Human; transporter protein; pharmacogenomic analysis; immune response;
KW	transgenic animal; gene therapy; chromosome 4; gene; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FH	variation
FT	Location/Qualifiers
FT	replace(1275, C)
FT	/*tag= a
FT	/standard name= "Single nucleotide polymorphism (SNP)"
FT	replace(1456, C)
FT	/*tag= b
FT	/standard name= "Single nucleotide polymorphism (SNP)"
FT	2159..62365
FT	/*tag= c
FT	/product= "Human protein"
FT	2159..2237
FT	/*tag= d
FT	2238..22041
FT	/*tag= e
FT	replace(5893, A)
FT	/*tag= f
FT	/standard name= "Single nucleotide polymorphism (SNP)"
FT	replace(6226, G)
FT	/*tag= g
FT	/standard name= "Single nucleotide polymorphism (SNP)"
FT	replace(8666, G)
FT	/*tag= h
FT	/standard name= "Single nucleotide polymorphism (SNP)"
FT	replace(10397, T)
FT	/*tag= i
FT	/standard name= "Single nucleotide polymorphism (SNP)"
FT	replace(10621, -)
FT	/*tag= j
FT	/standard name= "Single nucleotide polymorphism (SNP)"
FT	replace(19651, AG)

```
FT FT /*tag= k
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(15891, -)
FT FT /*tag= l
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(20272, A)
FT FT /*tag= m
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(20412, A)
FT FT /*tag= n
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT 22042. .22199
FT FT /*tag= o
FT FT 22200. .30359
FT FT /*tag= p
FT FT replace(23340, G)
FT FT /*tag= q
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(29948, A)
FT FT /*tag= r
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT 30360. .30459
FT FT /*tag= s
FT FT 30460. .31475
FT FT /*tag= t
FT FT 31476. .31663
FT FT /*tag= u
FT FT 31664. .32964
FT FT /*tag= v
FT FT 32965. .33087
FT FT /*tag= w
FT FT 33088. .34548
FT FT /*tag= x
FT FT replace(33579, C)
FT FT /*tag= y
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT 34549. .34755
FT FT /*tag= z
FT FT 34756. .37975
FT FT /*tag= aa
FT FT 37976. .38056
FT FT /*tag= ab
FT FT 38057. .39552
FT FT /*tag= ac
FT FT 39553. .40098
FT FT /*tag= ad
FT FT 40099. .46366
FT FT /*tag= ae
FT FT replace(40762, A)
FT FT /*tag= af
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(40936, C)
FT FT /*tag= ag
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(45998, G)
FT FT /*tag= ah
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT 46367. .46553
FT FT /*tag= ai
FT FT 46554. .49237
FT FT /*tag= aj
FT FT replace(47771, C)
FT FT /*tag= ak
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(48117, T)
FT FT /*tag= al
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT 49238. .49636
FT FT /*tag= am
FT FT 49637. .55445
FT FT /*tag= an
FT FT replace(54563, G)
FT FT /*tag= ao

FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT 55446. .55662
FT FT /*tag= ap
FT FT 55663. .62274
FT FT /*tag= aq
FT FT replace(58735, T)
FT FT /*tag= ar
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(59643, A)
FT FT /*tag= as
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(61638, G)
FT FT /*tag= at
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT 62275. .62362
FT FT /*tag= au
FT FT replace(63291, G)
FT FT /*tag= av
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(63463, G)
FT FT /*tag= aw
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(63636, A)
FT FT /*tag= ax
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT replace(63998, C)
FT FT /*tag= ay
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"

FT FT WO200272764-A2.
FT FT 19-SEP-2002.
FT FT 08-MAR-2002; 2002WO-US007156.
FT FT 13-MAR-2001; 2001US-00804472.
FT FT (PEKE ) PE CORP NY.
FT FT Shao W, Merkulov GV, Difrancesco V, Beasley EM;
FT FT WPI; 2002-759843/82.
FT FT DR P-PsDB; AAE29091.
FT FT XX
FT FT PT New human transporter peptides and polynucleotides encoding the peptides,
FT FT useful as models for developing human therapeutic targets, in gene
FT FT therapy, or for producing or eliciting an immune response.
FT FT PS Claim 4; Page 93-111, 114p; English.
FT FT XX
FT FT CC The invention relates to human transporter polypeptides and nucleic acid
FT FT molecules encoding such polypeptides. Sequences of the invention are
FT FT useful as models for the development of human therapeutic targets, in the
FT FT identification of therapeutic proteins and serve as targets for the
FT FT development of human therapeutic agents that modulate transporter
FT FT activity in cells and tissues that express the transporter. Polypeptides
FT FT of the invention can be used to elicit immune response, as reagents in
FT FT assays designed to determine the levels of the protein in biological
FT FT fluids, as markers for tissue in which the corresponding protein is
FT FT preferentially expressed, in the identification of modulators of the
FT FT peptides and in pharmacogenomic analysis. Nucleic acid molecules of the
FT FT invention are useful as hybridisation probes, in constructing vectors,
FT FT host cells or transgenic animals expressing all or a part of the nucleic
FT FT acid, for monitoring the effectiveness of modulating compounds on the
FT FT expression or activity of the transporter gene in clinical trials or in
FT FT treatment regimen, in gene therapy and as antisense constructs to control
FT FT CC transporter gene expression in cells, tissue and organisms. The present
FT FT CC sequence is human transporter gene located on chromosome 4
FT FT XX
FT FT SQ Sequence 65359 BP; 18337 A; 10760 C; 11602 G; 21226 T; 0 U; 3434 Other;
FT FT Query Match 6.5%; Score 35.8; DB 6; Length 65359;
FT FT Best Local Similarity 49.2%; Pred. No. 23;
```

	Matches	94;	Conservative	0;	Mismatches	97;	Indels	0;	Gaps	0;
Qy	5	CGAGTGTAAATTCCTGATGTAGCTACATTAAATCTTTATTCATCAATCAATTAAGATCAGT	64							
Db	52323	CCAATTTTTTTTTTTTTTTGGTAGAGACAAGGCTTGCTATGTACCCAGCTGGCTTGAT	52264							
Qy	65	CTTGCGGTACCTCTACGGGCTCTCACCATGCATTCAGATATCTCTGTAGAGCGAT	124							
Db	52263	CTCCTGTGTGCACAACAACCTCCACCTGGCCTCCCATTTAAATAATTTTTTGATGCTA	52204							
Qy	125	GTTATGCAAGAGCCCATTAAGATGAGACAAATCTTAATGAACAAGGCTATGACTGTGAAA	184							
Db	52203	CTCATTCAGAGGCTTTTAATAAGAACATTACAAATTTTCATTTCTTCATTTTTTTAAA	52144							
Qy	185	AACCAATTTGTA	195							
Db	52143	AAATATTTTAA	52133							

Search completed: September 16, 2006, 14:56:01
 Job time : 518 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2006, 15:58:20 ; Search time 171 Seconds

(without alignments)
6072.897 Million cell updates/sec

Title: US-10-815-774-5

Perfect score: 555
Sequence: 1 ttggcagatgtaattccctga.....atgtatccagctgtgattt 555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:
1: /EMC Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:*
2: /EMC Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:*
3: /EMC Celerra_SIDS3/ptodata/2/ina/6 COMB.seq:*
4: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:*
5: /EMC Celerra_SIDS3/ptodata/2/ina/8 COMB.seq:*
6: /EMC Celerra_SIDS3/ptodata/2/ina/9 COMB.seq:*
7: /EMC Celerra_SIDS3/ptodata/2/ina/10 COMB.seq:*
8: /EMC Celerra_SIDS3/ptodata/2/ina/11 COMB.seq:*
9: /EMC Celerra_SIDS3/ptodata/2/ina/12 COMB.seq:*
10: /EMC Celerra_SIDS3/ptodata/2/ina/13 COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	555	100.0	555	US-09-727-769A-5	Sequence 5, Appl1
2	555	100.0	1080	US-09-727-769A-7	Sequence 7, Appl1
3	344.2	62.0	555	US-09-324-910-5	Sequence 5, Appl1
4	344.2	62.0	555	US-09-793-495-5	Sequence 5, Appl1
5	344.2	62.0	1080	US-09-324-910-11	Sequence 11, Appl1
6	344.2	62.0	1080	US-09-793-495-11	Sequence 11, Appl1
7	36.4	6.6	474	US-09-621-976-18033	Sequence 18033, A
8	35.8	6.5	101894	US-09-949-016-12005	Sequence 12005, A
9	35.8	6.5	103894	US-09-949-016-14450	Sequence 14450, A
10	35.6	6.4	2489	US-09-141-047-7	Sequence 7, Appl1
11	34.8	6.3	601	US-09-949-016-23464	Sequence 23464, A
12	34.8	6.3	601	US-09-949-016-94061	Sequence 94061, A
13	34.8	6.3	1751	US-09-604-393B-3	Sequence 3, Appl1
14	34.8	6.3	50000	US-09-662-254B-25	Sequence 25, Appl1
15	34.6	6.2	87470	US-09-949-016-15881	Sequence 15881, A
16	34.6	6.2	1072	US-08-956-171E-435	Sequence 435, App
17	34.4	6.2	1072	US-08-781-986A-435	Sequence 435, App
18	34.4	6.2	3510	US-09-487-558B-105	Sequence 105, App
19	34.2	6.2	2019	US-10-009-966C-3	Sequence 3, Appl1
20	34.2	6.2	3666	US-08-682-517-13	Sequence 13, Appl1
21	34.2	6.2	3666	US-08-682-517-14	Sequence 14, Appl1
22	34.2	6.2	4197	US-08-682-517-7	Sequence 7, Appl1
23	34.2	6.2	4197	US-08-682-517-8	Sequence 8, Appl1

C 24	34	6.1	2286	3	US-10-104-047-1871	Sequence 1871, Ap
C 25	34	6.1	56694	3	US-09-949-016-12568	Sequence 12568, A
C 26	34	6.1	56702	3	US-09-949-016-15423	Sequence 15423, A
C 27	33.8	6.1	690	3	US-09-489-039A-6025	Sequence 6025, Ap
C 28	33.6	6.1	1141	3	US-09-806-708B-22	Sequence 22, Appl
C 29	33.6	6.1	4223	3	US-09-541-782-5	Sequence 5, Appl1
C 30	33.6	6.1	4223	3	US-09-723-820-5	Sequence 5, Appl1
C 31	33.6	6.1	4223	3	US-10-270-085-5	Sequence 5, Appl1
C 32	33.6	6.1	36791	3	US-09-949-016-17451	Sequence 17451, A
C 33	33.6	6.1	450395	3	US-09-949-016-15473	Sequence 15473, A
C 34	33.4	6.0	205163	3	US-09-949-016-17009	Sequence 17009, A
C 35	33.2	6.0	601	3	US-09-949-016-96083	Sequence 96083, A
C 36	33.2	6.0	601	3	US-09-949-016-96084	Sequence 96084, A
C 37	33.2	6.0	387902	3	US-09-949-016-14543	Sequence 14543, A
C 38	33.2	6.0	421883	3	US-09-949-016-12557	Sequence 12557, A
C 39	33.2	6.0	1664976	3	US-08-916-421B-1	Sequence 1, Appl1
C 40	33.2	6.0	1664976	3	US-09-692-570-1	Sequence 1, Appl1
C 41	33	5.9	37282	3	US-09-949-016-12440	Sequence 12440, A
C 42	33	5.9	37308	3	US-09-949-016-14162	Sequence 14162, A
C 43	33	5.9	37308	3	US-09-949-016-14163	Sequence 14163, A
C 44	33	5.9	218940	3	US-09-949-016-17539	Sequence 17539, A
C 45	32.8	5.9	601	3	US-09-949-016-35281	Sequence 35281, A

ALIGNMENTS

RESULT 1
US-09-727-769A-5
; Sequence 5, Application US/09727769A
; Patent No. 6756221
; GENERAL INFORMATION:
; APPLICANT: Amaro Enzyme, Inc.
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SA
; FILE REFERENCE: Q62106
; CURRENT APPLICATION NUMBER: US/09/727,769A
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: JP 11-345044
; PRIOR FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 555
; TYPE: DNA
; ORGANISM: *Cryoseobacterium* sp. No. 6756221 9670
US-09-727-769A-5

Query Match 100.0%; Score 555; DB 3; Length 555;
Best Local Similarity 100.0%; Pred. No. 2.1e-159;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTGGCAGTGTAAATTCCTGATGTACCTAATTAAATCTTTATTCATCAATAAAGAT	60
DB	1	TTGGCAGTGTAAATTCCTGATGTACCTAATTAAATCTTTATTCATCAATAAAGAT	60
QY	61	CAGTCTTGCGTACTTCAAGCGGCTTCAAGTCAATGACCACTTCAAGTCAATGAC	120
DB	61	CAGTCTTGCGTACTTCAAGCGGCTTCAAGTCAATGACCACTTCAAGTCAATGAC	120
QY	121	GGATGTTATGCAAGGCCCATTAAGATGAGCAAACTTAATGACCAAGCGCTATGCT	180
DB	121	GGATGTTATGCAAGGCCCATTAAGATGAGCAAACTTAATGACCAAGCGCTATGCT	180
QY	181	GAATTTATGCAAGGCCCATTAAGATGAGCAAACTTAATGACCAAGCGCTATGCT	240
DB	181	GAATTTATGCAAGGCCCATTAAGATGAGCAAACTTAATGACCAAGCGCTATGCT	240
QY	241	AGCTTCAAGCGTGTCAATTTGGTATGATTAATTAATGCTCCGAGTAAAGGAAAGGA	300
DB	241	AGCTTCAAGCGTGTCAATTTGGTATGATTAATTAATGCTCCGAGTAAAGGAAAGGA	300
QY	301	ATTATGATCCTTCACTATTTTCAAGCGGCTCTGTAAGATGACGATGAGAAACGCT	360

Db 301 ATTATGATCTCTCACTAATTTTCAAGCGCTCTGTACAGATACAGATGAGAAACGCT 360
Qy 361 TGGCTTAACACCTCTTGGCGATCTGCATCCGTTCTCTTATGCTAATACTGAGAAAT 420
Db 361 TGGCTTAACACCTCTTGGCGATCTGCATCCGTTCTCTTATGCTAATACTGAGAAAT 420
Qy 421 GTTTATTAACAGAGTCTAGTAATTTCTTACCTGTATGACAACTGATCAATACCAAC 480
Db 421 GTTTATTAACAGAGTCTAGTAATTTCTTACCTGTATGACAACTGATCAATACCAAC 480
Qy 481 TGTGTAAGTCACTAATTTTCACTGCTTCCGATGTTCTCTTACCTGACCTGACCGAGTGA 540
Db 481 TGTGTAAGTCACTAATTTTCACTGCTTCCGATGTTCTCTTACCTGACCTGACCGAGTGA 540
Qy 541 TCCAGCTGTGATTT 555
Db 541 TCCAGCTGTGATTT 555

RESULT 2

US-09-727-769A-7
; Sequence 7, Application US/09727769A
; Patent No. 6756221
; GENERAL INFORMATION:
; APPLICANT: Amamo Enzyme, Inc.
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SAME
; FILE REFERENCE: 062106
; CURRENT APPLICATION NUMBER: US/09/727,769A
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: JP 11-345044
; PRIOR FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Cryseobacterium sp. No. 6756221 9670
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1020)
; OTHER INFORMATION:
; NAME/KEY: mat peptide
; LOCATION: (466)..()
; OTHER INFORMATION:
; US-09-727-769A-7

Query Match

Best Local Similarity 100.0%; Score 555; DB 3; Length 1080;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCGAGTGAATTCCTGATGTAGTACATTAATTTCTTATTCATCAATAAAGAT 60
Db 466 TTGGCGAGTGAATTCCTGATGTAGTACATTAATTTCTTATTCATCAATAAAGAT 525
Qy 61 CAGTCTGGCGTACCTCTACGGCGTCTCACCATGATCACATTCAGATATCTGTAGAC 120
Db 526 CAGTCTGGCGTACCTCTACGGCGTCTCACCATGATCACATTCAGATATCTGTAGAC 585
Qy 121 GGAGTGAATGCAAGGCCATTAAGATGAGACAAATCTTAATGAACAACGGCTATGATGT 180
Db 586 GGAGTGAATGCAAGGCCATTAAGATGAGACAAATCTTAATGAACAACGGCTATGATGT 645
Qy 181 GAAAAACAATTTGTATACGAAACCTAAAGGATCAACAGGAATCTGTGTGGCGTGG 240
Db 646 GAAAAACAATTTGTATACGAAACCTAAAGGATCAACAGGAATCTGTGTGGCGTGG 705
Qy 241 AGTACCAAGTTCGATATTTGTATGCTATTAATAATGCTTCGGAATGACGAAAAAGA 300
Db 706 AGTACCAAGTTCGATATTTGTATGCTATTAATAATGCTTCGGAATGACGAAAAAGA 765
Qy 301 ATTATGATCTCTCACTAATTTTCAAGCGCTCTGTACAGATACAGATGAGAAACGCT 360

Db 766 ATTATGATCTCTCACTAATTTTCAAGCGCTCTGTACAGATACAGATGAGAAACGCT 825
Qy 361 TGGCTTAACACCTCTTGGCGATCTGCATCCGTTCTCTTATGCTAATACTGAGAAAT 420
Db 826 TGGCTTAACACCTCTTGGCGATCTGCATCCGTTCTCTTATGCTAATACTGAGAAAT 885
Qy 421 GTTTATTAACAGAGTCTAGTAATTTCTTACCTGTATGACAACTGATCAATACCAAC 480
Db 886 GTTTATTAACAGAGTCTAGTAATTTCTTACCTGTATGACAACTGATCAATACCAAC 945
Qy 481 TGTGTAAGTCACTAATTTTCACTGCTTCCGATGTTCTCTTACCTGACCTGACCGAGTGA 540
Db 946 TGTGTAAGTCACTAATTTTCACTGCTTCCGATGTTCTCTTACCTGACCTGACCGAGTGA 1005
Qy 541 TCCAGCTGTGATTT 555
Db 1006 TCCAGCTGTGATTT 1020

RESULT 3

US-09-324-910-5
; Sequence 5, Application US/09324910
; Patent No. 6251651
; GENERAL INFORMATION:
; APPLICANT: Matsura, Akira
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE
; FILE REFERENCE: Q-54362
; CURRENT APPLICATION NUMBER: US/09/324,910
; PRIOR FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: HEI-10-173940
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Cryseobacterium gleum
; US-09-324-910-5

Query Match

Best Local Similarity 62.0%; Score 344.2; DB 3; Length 555;
Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 7 AGTGAATTCCTGATGTAGTACATTAATTTCTTATTCATCAATAAAGATCACTCT 66
Db 7 AGTGAATTCCTGATGTAGTACATTAATTTCTTATTCATCAATAAAGATCACTCT 66
Qy 67 TGGGTAACCTTACAGACATCTTCTCTGTATCACTTACAGATCCGTTGACGATGT 126
Db 67 TGGGTAACCTTACAGACATCTTCTCTGTATCACTTACAGATCCGTTGACGATGT 126
Qy 127 TATGCAAGGCCATTAAGATGAGACAAATCTTAATGAACAACGGCTATGATGAAAA 186
Db 127 TATGCAAGGCCATTAAGATGAGACAAATCTTAATGAACAACGGCTATGATGAAAA 186
Qy 187 CAATTTGTATACGAAACCTAAAGCATCAACAGAACTTGTGTGTGGCTGAGCTAC 246
Db 187 CAATTTGTATACGAAACCTAAAGCATCAACAGAACTTGTGTGTGGCTGAGCTAC 246
Qy 247 CAGTTGCAATATTTGTATGCTTAATAATGCTTCCGAGTAAACGAAAAAGATTAAT 306
Db 247 CAGTTGCAATATTTGTATGCTTAATAATGCTTCCGAGTAAACGAAAAAGATTAAT 306
Qy 307 GATCTTCACTAATTTTCAAGCGGCTCTGTACAGATACAGATGAGAAACGCTTGCCTT 366
Db 307 GATCTTCACTAATTTTCAAGCGGCTCTGTACAGATACAGATGAGAAACGCTTGCCTT 366
Qy 367 AACACCTCTGCGGATCTGCATCCGTTCTCTTATGCTAATACTGAGAAATGTTTAT 426
Db 367 AACACCTCTGCGGATCTGCATCCGTTCTCTTATGCTAATACTGAGAAATGTTTAT 426


```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23364
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-23364

Query Match
Best Local Similarity 50.0%; Pred. No. 1.7;
Matches 87; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 22 GTAGCTACATTAATTTCTTATTCATCAATAAAGATCAGTCTGGCTACTCTAG 81
DB 592 GTAGAGACAAAGGTCTTGATGTACCCAGCTGGTCTTGATCTCCTGGCCAAACAA 533
QY 82 GCGTCTCAGCATGATCAGATATCTGTAGACGATGTATGCAAGGCCAT 141
DB 532 ACCCTCCACCTGGCTCCCATTTAAANATTTTGATGCTACTCATTCAGGCGCTT 473
QY 142 AAGATGAGCAATCTTAATGAACAACGGCTATGACTGTGAAAACAATTTGA 195
DB 472 TAAATAGGAACATTACATTTCTTCATTTTAAAAAATATTTTA 419

RESULT 12
US-09-949-016-94061/c
; Sequence 94061, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94061
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-94061

Query Match
Best Local Similarity 6.3%; Score 34.8; DB 3; Length 601;
Matches 87; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 22 GTAGCTACATTAATTTCTTATTCATCAATAAAGATCAGTCTGGCTACTCTAG 81
DB 592 GTAGAGACAAAGGTCTTGATGTACCCAGCTGGTCTTGATCTCCTGGCCAAACAA 533
QY 82 GCGTCTCAGCATGATCAGATATCTGTAGACGATGTATGCAAGGCCAT 141
DB 532 ACCCTCCACCTGGCTCCCATTTAAANATTTTGATGCTACTCATTCAGGCGCTT 473
QY 142 AAGATGAGCAATCTTAATGAACAACGGCTATGACTGTGAAAACAATTTGA 195
DB 472 TAAATAGGAACATTACATTTCTTCATTTTAAAAAATATTTTA 419
```

```

RESULT 13
US-09-004-393B-3
; Sequence 3, Application US/09004393B
; Patent No. 6310271
; GENERAL INFORMATION:
; APPLICANT: Hanson D., Andrew
; APPLICANT: Rathinasabapathi, Bala
; APPLICANT: Burnett, Michael
; TITLE OF INVENTION: Polynucleotides Encoding Choline Monooxygenase and
; TITLE OF INVENTION: Plants Transformed Therewith
; FILE REFERENCE: UP-162
; CURRENT APPLICATION NUMBER: US/09/004,393B
; CURRENT FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/035,147
; PRIOR FILING DATE: 1997-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1751
; TYPE: DNA
; ORGANISM: Beta vulgaris
US-09-004-393B-3

Query Match
Best Local Similarity 6.3%; Score 34.8; DB 3; Length 1751;
Matches 99; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 274 AATGCTTCGGAGTACGGAAGAAAGATTAATGATCTTCACTATTTTCAACGGCTCT 333
DB 643 AAGCCGACGAAAGCACTGAAAACAAAATTGGATCTTAAAGAACTTGGCGCTGCACCC 702
QY 334 GTPACAGATACAGCATGAGAAACGCTTGCCTTGGCGATCTGCATCCGT 393
DB 703 CTAAAGTTGCAAGATGGGCCCCATTCATTTATCAAGTTGGACCGATCTTGAATGCT 762
QY 394 TCCTCTTATGCTAATACCTCAGAAATGTTTATTAACAAGATCTAGTAATTTTAACTTG 453
DB 763 AATGCTGATGTTGGAACAGAGTGATGTTAATCTGCAAGATGTAAAGGCCCATGCT 822
QY 454 TATGACAAATCTGATCAATACCA 479
DB 823 TTTGATCTTAATCTAAAGTTCACCA 848

RESULT 14
US-09-662-254B-25
; Sequence 25, Application US/09662254B
; Patent No. 6931145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bowen, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; FILE REFERENCE: UP-221C1X1
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Version 3.1
; SEQ ID NO 25
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Amsacta moorei entomopoxvirus
US-09-662-254B-25

Query Match
Best Local Similarity 6.3%; Score 34.8; DB 3; Length 50000;
Matches 81; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

QY	116	TAGACGATGTTATGCAAGACCCATTAAGTGGACAAATCTTAAGAACCAAGCGCTAG	175
Db	28661	TTGGTGCATATTTTGGCTTTAAGGACCAATTAATACATTTAGCTAAGATTAATTCGAAT	28720
QY	176	ACTGTGAAAAACAATTGTGTATAGGAAACCTTAAGGCATCAACGAGACCTTGCTGTGG	235
Db	28721	TAAATGAAAAACATCTATTATTTCCATACATTAACATGTCACCTGTAAATTAATTTTGAAG	28780
QY	236	CGTGAAGCTACCAAGTTGCAATATTTGGTTAAGCTATAAA	273
Db	28781	GATGTAGTAACCCAGATGCTATATTTGATATATATAAAAA	28818

```

RESULT 15
US-09-949-016-15881/C
; Sequence 15881, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498-
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15881
; LENGTH: 87470
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (87470)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15881

```

Query Match	Similarity	6.2%;	Score 34.6;	DB 3;	Length 87470;
Best Local	Similarity	61.8%;	Pred. No.15;		
Matches	55;	Conservative	0;	Mismatches	34;
				Indels	0;
				Gaps	0;
QY	108	ATATCCGTGAGCGGATGTTATGCAAGGCCATTAAGTGAACAATCTTAATGAACA	167		
Db	43412	ATATCCCACTGATGCGCATGTGGCTACTGAGCACTTAAGAAATGGCTTAGTCTTAACAT	43353		
QY	168	CGGCTATGACTGTGAAATAACATTTGAT	196		
Db	43352	GTGCTGCAATGTATAAATTAATACTGAT	43324		

Search completed: September 16, 2006, 17:06:13
Job time : 177 secs

This Page Blank (uspto)


```
Db 121 GAGTGTATGACAGACCCCTAAGATGAGCAAAATCTTAATGAACAAGGCTATGACTT 180
Qy 181 GAAAAACAATTTGTATACGGAACCTAAGGACATCAACAGAACTTGCTGTGTGGCGTGG 240
Db 181 GAAAAACAATTTGTATACGGAACCTAAGGACATCAACAGAACTTGCTGTGTGGCGTGG 240
Qy 241 AGCTACCAAGTTCGAATATTTGTAGCTATAAAAATGCTTCGGAGTAAAGAAAAA 300
Db 241 AGCTACCAAGTTCGAATATTTGTAGCTATAAAAATGCTTCGGAGTAAAGAAAAA 300
Qy 301 ATTATGATCCTTCACATATTTTCAAGGCGCTGTACAGATACAGATAGAAACGCT 360
Db 301 ATTATGATCCTTCACATATTTTCAAGGCGCTGTACAGATACAGATAGAAACGCT 360
Qy 361 TGGGTAAACACCTTCGCGATCTGCATCCGTTTCTCTTATGCTAATATCGAGAAAT 420
Db 361 TGGGTAAACACCTTCGCGATCTGCATCCGTTTCTCTTATGCTAATATCGAGAAAT 420
Qy 421 GTTTATTACAGAACTCTAGTAATCTTACCTGTATGACAAACATCTGATCAATACCAAC 480
Db 421 GTTTATTACAGAACTCTAGTAATCTTACCTGTATGACAAACATCTGATCAATACCAAC 480
Qy 481 TGTCTACTGACTAAATTTTCACTGCTTCGCGATGTTCTGCTTCACTGACCGAGATGA 540
Db 481 TGTCTACTGACTAAATTTTCACTGCTTCGCGATGTTCTGCTTCACTGACCGAGATGA 540
Qy 541 TCCAGCTGTGATTT 555
Db 541 TCCAGCTGTGATTT 555
```

```
RESULT 2
US-10-815-774-7
; Sequence 7, Application US/10815774
; Publication No. US20040175799A1
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SAME
; FILE REFERENCE: 080844
; CURRENT APPLICATION NUMBER: US/10/815,774
; PRIOR FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US/09/727,769
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US/09/324,910
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: JP Hei. 11-345044
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP Hei. 10-173940
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Cryoseobacterium sp. No. 9670
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1020)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (466)..()
US-10-815-774-7
```

```
Query Match 100.0%; Score 555; DB 9; Length 1080;
Best Local Similarity 100.0%; Pred. No. 1.6e-153;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 TTGGCGAGTGTATTCCTGATGTAGCTACATTAATTTCTTATTCATCAATAAAGAT 60
Db 466 TTGGCGAGTGTATTCCTGATGTAGCTACATTAATTTCTTATTCATCAATAAAGAT 525
```

```
Qy 61 CAGTCTTCGGTACTCTTACAGCGGCTCTCAACATGATCACTTCAAGATATCTGTAGAC 120
Db 526 CAGTCTTCGGTACTCTTACAGCGGCTCTCAACATGATCACTTCAAGATATCTGTAGAC 585
Qy 121 GAGTGTATGCAAGAGCCCATTAAGATGAGCAAAATCTTAATGAACACGGCTAAGACTGT 180
Db 586 GAGTGTATGCAAGAGCCCATTAAGATGAGCAAAATCTTAATGAACACGGCTAAGACTGT 645
Qy 181 GAAAAACAATTTGTATACGGAACCTAAGGACATCAACAGAACTTGCTGTGTGGCGTGG 240
Db 646 GAAAAACAATTTGTATACGGAACCTAAGGACATCAACAGAACTTGCTGTGTGGCGTGG 705
Qy 241 AGCTACCAAGTTCGAATATTTGTAGCTATAAAAATGCTTCGGAGTAAAGAAAAA 300
Db 706 AGCTACCAAGTTCGAATATTTGTAGCTATAAAAATGCTTCGGAGTAAAGAAAAA 300
Qy 301 ATTATGATCCTTCACATATTTTCAAGGCGCTGTGACAGATACAGATAGAAACGCT 360
Db 766 ATTATGATCCTTCACATATTTTCAAGGCGCTGTGACAGATACAGATAGAAACGCT 825
Qy 361 TGGGTAAACACCTTCGCGATCTGCATCCGTTTCTCTTATGCTAATATCGAGAAAT 420
Db 826 TGGGTAAACACCTTCGCGATCTGCATCCGTTTCTCTTATGCTAATATCGAGAAAT 885
Qy 421 GTTTATTACAGAACTCTAGTAATCTTACCTGTATGACAAACATCTGATCAATACCAAC 480
Db 886 GTTTATTACAGAACTCTAGTAATCTTACCTGTATGACAAACATCTGATCAATACCAAC 945
Qy 481 TGTCTACTGACTAAATTTTCACTGCTTCGCGATGTTCTGCTTCACTGACCGAGATGA 540
Db 946 TGTCTACTGACTAAATTTTCACTGCTTCGCGATGTTCTGCTTCACTGACCGAGATGA 1005
Qy 541 TCCAGCTGTGATTT 555
Db 1006 TCCAGCTGTGATTT 1020
```

```
RESULT 3
US-09-793-495-5
; Sequence 5, Application US/09793495
; Publication No. US20040072318A1
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE SAME
; FILE REFERENCE: 0-54362
; CURRENT APPLICATION NUMBER: US/09/793,495
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/324,910
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: HEI-10-173940
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Cryoseobacterium gleum
US-09-793-495-5
```

```
Query Match 62.0%; Score 344.2; DB 3; Length 555;
Best Local Similarity 76.7%; Pred. No. 4.1e-91;
Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
```

```
Qy 7 AGTGTATTCCTGATGTAGCTACATTAATTTCTTATTCATCAATAAAGATCACT 66
Db 7 AGTGTATTCCTGATGTAGCTACATTAATTTCTTATTCATCAATAAAGATCACT 66
Qy 67 TGGGTACCTCTACAGCGGCTCTCAACATGATCAATTCAGATATCTGTAGACGATGT 126
Db 67 TGGGTACCTCTACAGCGGCTCTCTTGTATACCTTCAGATATCTGTAGACGATGT 126
```


QY 127 TATGCAAGGCCATTAGATGAGCAAACTTTAATGACAACGGCTATGACTGTGAAA 186
| | | | |
DB 127 TATGCAAGGCCATTAGATGAGCAAACTTTAATGACAACGGCTATGACTGTGAAA 186
| | | | |
QY 187 CAATTGTATAGGAACTTAAAGGATCAACAGAACTTGTGTGTGGCGTGGAGCTAC 246
| | | | |
DB 187 CAGTTGTATAGGAACTTAAAGGATCAACAGAACTTGTGTGTGGCGTGGAGCTAC 246
| | | | |
QY 247 CAGTTGTATAGGAACTTAAAGGATCAACAGAACTTGTGTGTGGCGTGGAGCTAC 246
| | | | |
DB 247 CAGTTGTATAGGAACTTAAAGGATCAACAGAACTTGTGTGTGGCGTGGAGCTAC 246
| | | | |
QY 307 GATCTTCACTATTTTCAAGCGGCTCTGTAAACAGATGAGAAAGCGTTCCTT 366
| | | | |
DB 307 GATCTTCACTATTTTCAAGCGGCTCTGTAAACAGATGAGAAAGCGTTCCTT 366
| | | | |
QY 367 AACACTCTTGGGATCTGATCGCTTCTCTTATGCTAATCTGAGAAATGTTAT 426
| | | | |
DB 367 AACACTCTTGGGATCTGATCGCTTCTCTTATGCTAATCTGAGAAATGTTAT 426
| | | | |
QY 427 TACAGAGTCTAGTAATTTTACCTGTATGACAACATCTGATCAATCAACTGTGTA 486
| | | | |
DB 427 TACAGAGTCTAGTAATTTTACCTGTATGACAACATCTGATCAATCAACTGTGTA 486
| | | | |
QY 487 CTGACTAAATTTTCACTGCTTCCGAGATGTTCTCTTCACTGACCGGATGATCCAGC 546
| | | | |
DB 487 TTAACATATTTCTCATCCCTTTCAGAGATGTTCTCTTCCCGAGACCAAGTGTAGCAAC 546
| | | | |
QY 547 TGTGATTT 555
| | | | |
DB 547 TGTGATTT 555
| | | | |

RESULT 4

US-10-815-751-5
; Sequence 5, Application US/10815751
; Publication No. US20040166558A1
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; APPLICANT: Matsuzaki, Akira
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE
; FILE REFERENCE: Q-54362
; CURRENT APPLICATION NUMBER: US/10/815,751
; PRIOR FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US/09/793,495
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/324,910
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: HEI-10-173940
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Chryseobacterium gleum
US-10-815-751-5

Query Match 62.0%; Score 344.2; DB 8; Length 555;
Best Local Similarity 76.7%; Pred. No. 4.1e-91;
Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 7 AGTGTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 66
| | | | |
DB 7 AGTGTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 66
| | | | |
QY 67 TCGGATCTTCTAGCGGCTCTACACATGATCAACATGATGATGATGATGATGATGATGAT 126
| | | | |
DB 67 TCGGATCTTCTAGCGGCTCTACACATGATCAACATGATGATGATGATGATGATGATGAT 126
| | | | |
QY 127 TATGCAAGGCCATTAGATGAGCAAACTTTAATGACAACGGCTATGACTGTGAAA 186
| | | | |
DB 127 TATGCAAGGCCATTAGATGAGCAAACTTTAATGACAACGGCTATGACTGTGAAA 186
| | | | |

DB 127 TATGCAAGGCCATTAGATGAGCAAACTTTAATGACAACGGCTATGACTGTGAAA 186
| | | | |
QY 187 CAATTGTATAGGAACTTAAAGGATCAACAGAACTTGTGTGTGGCGTGGAGCTAC 246
| | | | |
DB 187 CAGTTGTATAGGAACTTAAAGGATCAACAGAACTTGTGTGTGGCGTGGAGCTAC 246
| | | | |
QY 247 CAGTTGTATAGGAACTTAAAGGATCAACAGAACTTGTGTGTGGCGTGGAGCTAC 246
| | | | |
DB 247 CAGTTGTATAGGAACTTAAAGGATCAACAGAACTTGTGTGTGGCGTGGAGCTAC 246
| | | | |
QY 307 GATCTTCACTATTTTCAAGCGGCTCTGTAAACAGATGAGAAAGCGTTCCTT 366
| | | | |
DB 307 GATCTTCACTATTTTCAAGCGGCTCTGTAAACAGATGAGAAAGCGTTCCTT 366
| | | | |
QY 367 AACACTCTTGGGATCTGATCGCTTCTCTTATGCTAATCTGAGAAATGTTAT 426
| | | | |
DB 367 AACACTCTTGGGATCTGATCGCTTCTCTTATGCTAATCTGAGAAATGTTAT 426
| | | | |
QY 427 TACAGAGTCTAGTAATTTTACCTGTATGACAACATCTGATCAATCAACTGTGTA 486
| | | | |
DB 427 TACAGAGTCTAGTAATTTTACCTGTATGACAACATCTGATCAATCAACTGTGTA 486
| | | | |
QY 487 CTGACTAAATTTTCACTGCTTCCGAGATGTTCTCTTCACTGACCGGATGATCCAGC 546
| | | | |
DB 487 TTAACATATTTCTCATCCCTTTCAGAGATGTTCTCTTCCCGAGACCAAGTGTAGCAAC 546
| | | | |
QY 547 TGTGATTT 555
| | | | |
DB 547 TGTGATTT 555
| | | | |

RESULT 5

US-09-793-495-11
; Sequence 11, Application US/09793495
; Publication No. US20040072318A1
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; APPLICANT: Matsuzaki, Akira
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE
; FILE REFERENCE: Q-54362
; CURRENT APPLICATION NUMBER: US/09/793,495
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/324,910
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: HEI-10-173940
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Chryseobacterium gleum
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (61)..(462)
; NAME/KEY: mat_peptide
; LOCATION: (463)..(1017)
; NAME/KEY: CDS
; LOCATION: (61)..(1017)
US-09-793-495-11

Query Match 62.0%; Score 344.2; DB 3; Length 1080;
Best Local Similarity 76.7%; Pred. No. 5.7e-91;
Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 7 AGTGTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 66
| | | | |
DB 469 AGTGTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528
| | | | |
QY 67 TCGGATCTTCTAGCGGCTCTACACATGATCAACATGATGATGATGATGATGATGATGAT 126
| | | | |
DB 529 TCGGATCTTCTAGCGGCTCTACACATGATCAACATGATGATGATGATGATGATGATGAT 588
| | | | |

OY	127	TATGAAAGAGCCATPAAGATGAGCAATCTTAATGAACAACGGCTATGACTGGAATAA	186
Db	589	TATGCAAGGGCTCACAATAAGACAAATCTTATGAACGCCGCTATGACTGGAAGAAG	648
OY	187	CAATTTGTATACGGAACCTTAAGGCATCAAACGAACCTTCTGTGTGGCGTGAAGCTAC	246
Db	649	CAGTTGCTATATGTGAATCTGAGAGCTTCTACAGGAACATCTGTGTATCTGGGTATAT	708
OY	247	CACGTTGAATATTTGGTAAAGCTATAAATAATGCTTCGGAGTAAAGGAAAAAGAAATTAT	306
Db	709	CACGTAGCAATTTGTGTAAAGCTTCAAAATGCTTTCAGGAATTTGTTGAAAAAGAAATCAT	768
OY	307	GATCCTTCATAATTTTCAAGCGGTCCTGTAAACAGATACAGCATGAGAGAAACGTTGCGTT	366
Db	769	GATCCTTCATATATCTCCAGCGGTCCTGTAAACAGATCTGCATGAGAGAGCTCATGTACC	828
OY	367	AACACCTCTGGGATCTGCACTCGGTTCTCTTATATGTAATATCTGCAGGAATAGTTAT	426
Db	829	AACACAAAGCTGGAGATCTGCGTCTGTATCTTCTAAGCCAAATACAGCAAGGAAATGTTTAC	888
OY	427	TACAGAAGTCCAGAAATTTCTTACTCTGTATGACAACAATCTGATCAATATCAACTGTGTA	486
Db	889	TACAGAAATCCGTCAAGTTCTATTACTGTATATGATAACAATATGTGTAATACCAATGTGTA	948
OY	487	CTGACTAAATTTTCACTGCTTTCCGAGATGTTCTCTTCACTGCACCGGAGATGTATCCAGC	546
Db	949	TTAAACATATTTCTCATCCCTTTCAAGATGTTCTCTCTCCCAAGCAACAAAGTGTAGCAAGC	1008
OY	547	TGTGATTT 555	
Db	1009	TGTGGATTT 1017	

OY		7	TGCTATTCCTCATAGTACATTAATAATTCTTTATTCATCAATTAAGAATCGGT	66
Db		469	AGGTATATCTCCTACTCTGGACAACGCTGAACAGTTTATTACCAGATCAAAAACAGGT	528
OY		67	TGCAGTACCTCTACGGCGTCCCTCACCATGATCACAATTCAGATATCTGTAAACGATGT	126
Db		529	TGGGAACCTCTACAGCATCTTCTCCCTGTATCACCCTTGATATCCGGTTGACGGATGT	588
OY		127	TATCGAAGAGCCCATTAAGATGAGACAAATTTTATATGAACAACGGCTATATCATCTGTAAAA	186
Db		589	TATCAAAGGGCTCAAAATAGAGCAAAATCTATTAACGCCGGCTATGACTGTGAAAA	648
OY		187	CAATTTGATACGGAACCTAAAGGCATCAACAGAACTGTGTGTGGCGTGAAGCTAC	246
Db		649	CAGTTTCGATATGTGTATCTGAGAGCTTCTACAGAAACATGCTGTATCATGGGTATAT	708
OY		247	CACGTTCCAAATATTGTGTACCTATAAAAATGCTTCCGAGTAAAGAAAAAGATTAAT	306
Db		709	CACGTACCAATTTTGTGTAGCTTCAAAAATGCTTCAAGAAATGTTTAAAAAAGAACATAT	768
OY		307	GATCCTTCACATTTTCAAGCGGTCCGTGAACAGATACAGATGAGAGAAACGTTGCGTT	366
Db		769	GATCCTTCATATTCTCCAGCGGTCCGTGAACAGATTCGCAAGAGAGCTGATGACC	828
OY		367	AACACCTTCGCGGATCGATCGATCGTTTTCTCTTATGCTAATCTCAGAGAAATGTTAT	426
Db		829	AACCAAGCTGGGATCTGGGTCTGTATCTTCTTAGCCCAATACAGAGAAATGTTAC	888
OY		427	TACAGAACTCTAGTAATCTTACTGTATGACAACAATCTGATCAATACCACTGTGTA	486
Db		889	TACGAAGTCCGTAGGTTCAATTACTGATGATTAACAACAATGGAATACCAATTGTGTA	948
OY		487	CTGACTAAATTTTACTGCTTTCCGATGTTTCTTCACTGACCTGACCGGATGATCCAGC	546
Db		949	TTAAACAATATCTCATCCCTTTGAGATGTTCTCTTCCCACAGACCAAGTGTAGAACGC	1008
OY		547	TGTGATTTT	555
Db		1009	TGTGATTT	1017
RESULT 7				
US-10-723-860-5703				
Sequence 5703, Application US/10723860				
Publication No. US20040253606A1				
GENERAL INFORMATION:				
APPLICANT: Aziz, Matasha				
APPLICANT: Ginsburg, Wendy M.				
APPLICANT: Zlonick, Albert				
TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &				
TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators				
FILE REFERENCE: 05882, 0193, NEU001				
CURRENT APPLICATION NUMBER: US/10/723, 860				
PRIOR FILING DATE: 2003-11-26				
PRIOR APPLICATION NUMBER: 60/429, 739				
PRIOR FILING DATE: 2002-11-26				
NUMBER OF SEQ ID NOS: 8393				
SOFTWARE: PatentIn version 3.2				
SEQ ID NO 5703				
LENGTH: 1105				
TYPE: DNA				
ORGANISM: Homo sapiens				
US-10-723-860-5703				

```

; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: (61)..(1017)
US-10-815-751-11

Query Match      62.0%;      Score 344.2; DB 8;      Length 1080;
Best Local Similarity 76.7%;      Pred. No. 5,7e-91;
Matches 421; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
```

Query Match	7.0%	Score 39	DB 9	Length 1105
Best Local Similarity	52.8%	Pred. No. 2.8		
Matches 84	Conservative 0	Mismatches 75	Indels 0	Gaps 0

QY	182	AAAAACAATTGTATCGGAAACCTTAAAGCATCAACAGGAAGCTTCGTGCGCGCTGGA	2411
Db	91	AAAAGATGTTTATGGAAGAAATCTGGTTACTCTTGATATTGGAACCTTTATGAGATTTAAA	150
QY	242	GCTACACAGTGCATATTGTTAGCTATTAAAAATGCTTCGAGTGTACGAAAAAAGAA	301

Db 151 GCTACCTGTTATATAGGGGTTAGGCTAGAGAGCCCAAGAAAAAGGTTAAATTTG 210
Qy 302 TTTATGATCTTCACTATTTTTCAGGCGTCTGTACAG 340
Db 211 CTAAAGAAATTCACAAATTTAAAGTTGCTTTAAAAAG 249

RESULT 8

US-09-925-065A-874704
; Sequence 874704, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 874704
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-874704

Query Match Best Local Similarity 6.8%; Score 37.6; DB 4; Length 622;

Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 146 TGAGCAAACTTAAATGAACAACGGCTATGACTGTGAAAAACAATTTGTATCGGAACC 205
Db 462 TGAATATATGTTAAATTAATTAATGCTATTAATATGAACACCATTAATATAAAAAAT 521
Qy 206 TAAAGCATCAACAGAACTTCTGTGTGCGGTGAGCTA 245
Db 522 TAAAAAGAGCAACAGAAATTTGGAGATGCTTTGTAACCTA 561

RESULT 9

US-09-925-065A-874705
; Sequence 874705, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 874705
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-874705

Query Match Best Local Similarity 6.8%; Score 37.6; DB 4; Length 622;

Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 146 TGAGCAAACTTAAATGAACAACGGCTATGACTGTGAAAAACAATTTGTATCGGAACC 205
Db 462 TGAATATATGTTAAATTAATTAATGCTATTAATATGAACACCATTAATATAAAAAAT 521
Qy 206 TAAAGCATCAACAGAACTTCTGTGTGCGGTGAGCTA 245
Db 522 TAAAAAGAGCAACAGAAATTTGGAGATGCTTTGTAACCTA 561

RESULT 10

US-09-925-065A-874706
; Sequence 874706, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 874706
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-874706

Query Match Best Local Similarity 6.8%; Score 37.6; DB 4; Length 622;

Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 146 TGAGCAAACTTAAATGAACAACGGCTATGACTGTGAAAAACAATTTGTATCGGAACC 205
Db 462 TGAATATATGTTAAATTAATTAATGCTATTAATATGAACACCATTAATATAAAAAAT 521
Qy 206 TAAAGCATCAACAGAACTTCTGTGTGCGGTGAGCTA 245
Db 522 TAAAAAGAGCAACAGAAATTTGGAGATGCTTTGTAACCTA 561

RESULT 11

US-09-925-065A-874707
; Sequence 874707, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096

```

; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 874707
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-874707
```

```

Query Match      6.8%; Score 37.6; DB 4; Length 622;
Best Local Similarity 61.0%; Pred. No. 4.3;
Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```

QY      146 TGAGACAAATCTTAATGAACAAGCGCTATGCTGTGAAAAACAATTGTATACGGAACC 205
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      462 TGAATATATCTTAATTAATTAATTAATGCGTATTAATATGAAACACCATTAATAAAAAAAT 521
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      206 TAAAGCATCAACAGAACTTGCTGTGTGGCGTGAAGCTA 245
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      522 TAAAAAGCAACCAAGAAATATTGGAGATGCTTTGTAACCTA 561
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 12

```

US-09-925-065A-874704
; Sequence 874704, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 874704
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-874704
```

```

Query Match      6.8%; Score 37.6; DB 5; Length 622;
Best Local Similarity 61.0%; Pred. No. 4.3;
Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```

QY      146 TGAGACAAATCTTAATGAACAAGCGCTATGCTGTGAAAAACAATTGTATACGGAACC 205
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      462 TGAATATATCTTAATTAATTAATTAATGCGTATTAATATGAAACACCATTAATAAAAAAAT 521
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      206 TAAAGCATCAACAGAACTTGCTGTGTGGCGTGAAGCTA 245
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      522 TAAAAAGCAACCAAGAAATATTGGAGATGCTTTGTAACCTA 561
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 13

```

US-09-925-065A-874705
; Sequence 874705, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 874705
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-874705
```

```

Query Match      6.8%; Score 37.6; DB 5; Length 622;
Best Local Similarity 61.0%; Pred. No. 4.3;
Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```

QY      146 TGAGACAAATCTTAATGAACAAGCGCTATGCTGTGAAAAACAATTGTATACGGAACC 205
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      462 TGAATATATCTTAATTAATTAATTAATGCGTATTAATATGAAACACCATTAATAAAAAAAT 521
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      206 TAAAGCATCAACAGAACTTGCTGTGTGGCGTGAAGCTA 245
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      522 TAAAAAGCAACCAAGAAATATTGGAGATGCTTTGTAACCTA 561
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 14

```

US-09-925-065A-874706
; Sequence 874706, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 874706
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-874706
```

```

Query Match      6.8%; Score 37.6; DB 5; Length 622;
Best Local Similarity 61.0%; Pred. No. 4.3;
Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```


This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 16, 2006, 18:31:40 / Search time 193 Seconds
(without alignments)
5043.212 Million cell updates/sec

Title: US-10-815-774-5

Perfect score: 555

Sequence: 1 ttcgcgagtgatcctcctga.....atgatccagctggtattc 555

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2330496 seqs, 876882855 residues.

Total number of hits satisfying chosen parameters: 4660992

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
10: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	39.8	7.2	634888	6	US-10-533-365-1	Sequence 1, Appl
2	34.4	6.2	777	6	US-10-471-571A-2813	Sequence 2813, Ap
3	34.4	6.2	3510	8	US-11-217-529-191047	Sequence 191047, A
4	34	6.1	822	8	US-11-266-748A-40049	Sequence 40049, A
5	34	6.1	822	8	US-11-266-748A-214137	Sequence 214137, A
6	34	6.1	1744	9	US-11-218-305-11243	Sequence 11243, A
7	34	6.1	4620	8	US-11-266-748A-27304	Sequence 27304, A
8	34	6.1	4620	8	US-11-266-748A-31292	Sequence 31292, A
9	33.6	6.1	349	7	US-11-314-834-716	Sequence 716, Appl
10	33.6	6.1	339234	6	US-10-669-920-1264	Sequence 1264, Appl
11	33.4	6.0	1000	8	US-11-266-748A-207141	Sequence 207141, A
12	33.4	6.0	4647455	6	US-10-641-321-205	Sequence 205, Appl
13	33.2	6.0	397	8	US-11-266-748A-411280	Sequence 411280, A
14	33	5.9	1000	8	US-11-266-748A-206397	Sequence 206397, A
15	33	5.9	1467	6	US-10-953-349-36486	Sequence 36486, A
16	33	5.9	1602	9	US-11-218-305-16632	Sequence 16632, A
17	33	5.9	2582	9	US-11-218-305-16633	Sequence 16633, A
18	33	5.9	100137	6	US-10-540-898-620	Sequence 620, Appl
19	33	5.9	150000	8	US-11-266-748A-23591	Sequence 23591, A
20	33	5.9	242825	8	US-11-266-748A-32745	Sequence 32745, A
21	32.8	5.9	612	8	US-11-266-748A-255476	Sequence 255476, A
22	32.8	5.9	612	8	US-11-266-748A-315993	Sequence 315993, A
23	32.8	5.9	4581	8	US-11-217-529-79437	Sequence 79437, A

24	32.8	5.9	5504	7 US-11-329-884-1	Sequence 1, Appl
25	32.8	5.9	6124	7 US-11-329-884-21	Sequence 21, Appl
26	32.8	5.9	6133	7 US-11-329-884-2	Sequence 2, Appl
27	32.8	5.9	6133	7 US-11-363-149-38	Sequence 38, Appl
28	32.8	5.9	6133	7 US-11-363-151-38	Sequence 38, Appl
29	32.8	5.9	6133	8 US-11-266-748A-29518	Sequence 29518, A
30	32.8	5.9	6133	8 US-11-266-748A-56391	Sequence 56391, A
31	32.8	5.9	6228	7 US-11-329-884-23	Sequence 23, Appl
32	32.6	5.9	201	6 US-10-284-444-30584	Sequence 30584, A
33	32.6	5.9	21909	8 US-11-266-748A-24012	Sequence 24012, A
34	32.4	5.8	555	7 US-11-292-078-2805	Sequence 2805, Ap
35	32.4	5.8	1000	8 US-11-266-748A-202893	Sequence 202893,
36	32.4	5.8	3715	8 US-11-266-748A-78432	Sequence 78432, A
37	32.4	5.8	3715	8 US-11-266-748A-110231	Sequence 110231,
38	32.4	5.8	3715	8 US-11-266-748A-131243	Sequence 131243,
39	32.4	5.8	70665	6 US-10-505-928-596	Sequence 596, App
40	32.4	5.8	523543	6 US-10-540-898-308	Sequence 308, App
41	32.2	5.8	4337	8 US-11-293-697-1232	Sequence 1232, Ap
42	32	5.8	1000	8 US-11-266-748A-199526	Sequence 199526,
43	32	5.8	1029	8 US-11-217-529-78011	Sequence 78011, A
44	31.8	5.7	2226	6 US-10-471-571A-1893	Sequence 1893, Ap
45	31.8	5.7	2676	8 US-11-216-545-1960	Sequence 1960, Ap

ALIGNMENTS

RESULT 1	US-10-533-365-1	Application US/10533365
Sequence 1, Appl	US-10-533-365-1	Application US/10533365
GENERAL INFORMATION:	US2006014162A1	
APPLICANT:	Reynolds, Inga	
APPLICANT:	Gulcher, Jeffrey R.	
APPLICANT:	Grant, Struan F.	
APPLICANT:	Thorleifsson, Gudmar	
TITLE OF INVENTION:	Human Type II Diabetes Gene - Silt-3	
TITLE OF INVENTION:	Located on Chromosome 5q35	
FILE REFERENCE:	2345.2046-007	
CURRENT APPLICATION NUMBER:	US/10/533,365	
CURRENT FILING DATE:	2005-04-29	
PRIOR APPLICATION NUMBER:	PCT/US03/34801	
PRIOR FILING DATE:	2003-10-31	
PRIOR APPLICATION NUMBER:	60/423,541	
PRIOR FILING DATE:	2002-11-01	
NUMBER OF SEQ ID NOS:	236	
SOFTWARE:	FASTSEQ for Windows Version 4.0	
SEQ ID NO 1		
LENGTH:	634888	
TYPE:	DNA	
ORGANISM:	homo sapiens	
US-10-533-365-1		
Query Match	7.2%; Score 39.8; DB 6; Length 634888;	
Best Local Similarity	49.3%; Pred. No. 0.61;	
Matches 104; Conservative	0; Mismatches 107; Indels 0; Gaps 0;	
18	TCATGAGCTAATTAATTTCTTATTCATGAATTAAGANTCAGTTCGGTACCC	77
385208	TGCTTTGTTTAACATCAATCTATACACAGTTCACATATGACATTTACTGTTAGCC	385267
78	TACGCGCTCCTACATGATCATCATTCATATCTGTGAGACGATGTTATGACAGCC	137
385268	TTTTTGTTCACAAACATGCTGCTTCATACAGAGCTCTGACACATGAGTGAAGC	385327
138	CCATAGATGACAAATCTTAATGACACGCTATGACTGTGAAAAAATTTGTATA	197
385328	ACATTAATATGCTAGTCCAAATTAAGTATGCTATGAATTAATAATCAAGTGAT	385387
198	CGGAACCTTAAGCATCAACGAACTGC	228
385388	TCCAAAACGATTTCAAAAAAAGAAATGC	385418

```
RESULT 2
US-10-471-571A-2813/c
; Sequence 2813, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWIn99, version 1.03
; SEQ ID NO 2813
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-2813

Query Match
Best Local Similarity 6.2%; Score 34.4; DB 6; Length 777;
Matches 56; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 329 GTCTGTAAAGATACGATGAGAAACGCTTACACCTCTTGGGATCTGCAT 388
DB 247 GTTCTGAATAATATATATATGAGTGGCTTACGCTTACACCTCTTGGTTTACAT 188

QY 389 CCGTTCTCTTATGCTAATATGCTAGAGAAAT 420
DB 187 CAATGTCATTTCACGTAATTTTTCAGTTACT 156

RESULT 3
US-11-217-529-191047
; Sequence 191047, Application US/11217529
; Publication No. US2006009961A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 191047
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-191047

Query Match
Best Local Similarity 6.2%; Score 34.4; DB 8; Length 3510;
Matches 80; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
```

```
DB 3268 ACAAGCAACGCTCTGAATAATAGGACAAATGTTTCT 3303

RESULT 4
US-11-266-748A-40049/c
; Sequence 40049, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 40049
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (806)..(806)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-40049

Query Match
Best Local Similarity 6.1%; Score 34; DB 8; Length 822;
Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 102 ATTCAATATCTCTGTAGACGATGTTATGCAAGAGCCATTAAGTGAACAATCTTAA 161
DB 552 ATTCATTTTAAACCGGAGAAATATACATAGCAAGGTAAACAGTGAGCCAGACCTAA 493

QY 162 GAACAACGCTATGACTGTGAATAACAATTGTATACGGAACCTTAAGGCATCAACAG 221
DB 492 CAATTAATGCTTACACCTGAATATATGTAAAGTCTTAAATGCCAATAGTACTAACA 433

QY 222 AACTTGCTGT 231
DB 432 GACTTGATAT 423

RESULT 5
US-11-266-748A-214137/c
; Sequence 214137, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
```



```

; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 214137
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (806)..(806)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-214137
```

```

Query Match          6.1%; Score 34; DB 8; Length 822;
Best Local Similarity 53.8%; Pred. No. 1.3;
Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
```

```

QY 102 ATTACATATTCCTGTAGACGATGTATATGCAAGCCCAATAATGACCAATCTTAT 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 552 ATTCAATTTTACCGGAGAAATACATACGACAGGTAACCTGTGAGCCAGCACTTAT 493
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 162 GAACAACGGCTATGACTGTGAAAAACAATTGTATACGAAACCTTAAGGCATCAACAG 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 492 CAATATGCTTACGACTGAATAATATGTAAAGTCTTAAATGCCAATGTACTAACAAG 433
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 222 AACTTGCTGT 231
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 432 GACTTGATAT 423
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 6
US-11-218-305-11243
; Sequence 11243, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaifd, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn.
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11243
; LENGTH: 1744
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-11243
```

```

Query Match          6.1%; Score 34; DB 9; Length 1744;
Best Local Similarity 51.3%; Pred. No. 2;
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

```

QY 278 CTTCCGAGTAACGAAAAAGAAATTATGATCCTTCACTATTTCACGCGTCTGTAA 337
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1183 CTGACAGAGAAATGACATGAATTAATCTGTGATGTACTCTTGAAATCAGATCAGAA 1242
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 338 CAGATACAGCATGAGAAACCTTGCTTAAACCTCTTGCGGATGTGATCGTTTCT 397
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1243 CAGGTTCAGATGAGCAACCTTATCTATATCTACCCCACTCTGCATCAGCAAG 1302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 398 CTTATGCTAATACGTGACGAAATGTTTATTAACG 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1303 CGTACTAATATGTTGCTGAGAGGTGACTCAAG 1336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 7
US-11-266-748A-27304/C
; Sequence 27304, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27304
; LENGTH: 4620
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-27304
```

```

Query Match          6.1%; Score 34; DB 8; Length 4620;
Best Local Similarity 53.8%; Pred. No. 3.4;
Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
```

```

QY 102 ATTACATATTCCTGTAGACGATGTATATGCAAGCCCAATAATGACCAATCTTAT 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2176 ATTCAATTTTACCGGAGAAATACATACGACAGGTAACGTGAGCCAGCACTTAT 2117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 162 GAACAACGGCTATGACTGTGAAAAACAATTGTATACGAAACCTTAAGGCATCAACAG 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2116 CAATATGCTTACGACTGAATAATATGTAAAGTCTTAAATGCCAATGTACTAACAAG 2057
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 222 AACTTGCTGT 231
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2056 GACTTGATAT 2047
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 8
US-11-266-748A-11292/C
; Sequence 11292, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
```

```
APPLICANT: Harkin, Paul
APPLICANT: Johnson, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
PRIOR FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 48396
SOFTWARE: PatentIn version 3.3
SEQ ID NO 31292
LENGTH: 4620
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-31292
```

```
Query Match
Best Local Similarity 53.8%; Pred. No. 3.4;
Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
```

```
QY 102 ATTGAGATATCTGTAGACGATGTTATGCAAGGCCCATTAAGTGAACAATCTTAAT 161
    |||||
DB 2176 ATTCAATTTTACCGGAGAAAATACATAGCAAGGATTAACAGTGAGCCAACTTAAT 2117
    |||||
QY 162 GAACACGGCTATGACCTGTGAAAAACAATTGTATACGAAACCTTAAGGATCAAGG 221
    |||||
DB 2116 CAATTAATGCTTACGCACTGAATTAATGTAAAGTCTTCAATATGCCAATAGTAACAAGA 2057
    |||||
QY 222 AACTGCTGT 231
    |||||
DB 2056 GACTTGATAT 2047
    |||||
```

```
RESULT 9
US-11-314-834-716
Sequence 716, Application US/11314834
Publication No. US20060195942A1
GENERAL INFORMATION:
APPLICANT: Fincher, Karen L.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 16517.360 - 38-21(51770)C/US
CURRENT APPLICATION NUMBER: US/11/314,834
CURRENT FILING DATE: 2005-12-22
PRIOR APPLICATION NUMBER: US 09/732,627
PRIOR FILING DATE: 2000-12-08
NUMBER OF SEQ ID NOS: 4930
SEQ ID NO 716
LENGTH: 349
TYPE: DNA
ORGANISM: Gosypium hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: LIB3493-050-PI-M1-C3
US-11-314-834-716
```

```
Query Match
Best Local Similarity 55.0%; Pred. No. 1.2; Length 349;
```

```
Matches 66; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 326 GCGTCTGTGAAGATACGATGAGAAACCTTCCGTACACCTCTGGGATCTG 385
    |||||
DB 22 GGGCTGGCTAAGAAATGCTTGAGCTTTGGCCCAAGAACCATCTCAGGCTGCTG 81
    |||||
QY 386 CATCGTTTCTCTTATGTAATATGTCAGAAATGTTATTAAGAAAGCTTAATAT 445
    |||||
DB 82 CATCTGCTATCTGTTCTTCTCAATGTTGAGAAAGTTCCCAAAATAGTGAATT 141
    |||||
```

```
RESULT 10
US-10-669-920-1264
Sequence 1264, Application US/10669920
Publication No. US20060194265A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Malandro, Marc S.
TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
FILE REFERENCE: 20366-066001
CURRENT APPLICATION NUMBER: US/10/669,920
PRIOR FILING DATE: 2003-09-23
PRIOR APPLICATION NUMBER: US 10/004,113
PRIOR FILING DATE: 2001-10-23
PRIOR APPLICATION NUMBER: US 10/052,482
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 09/997,722
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: US 10/034,650
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: US 10/085,117
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 10/087,192
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 10/322,281
PRIOR FILING DATE: 2002-12-17
PRIOR APPLICATION NUMBER: US 10/322,696
PRIOR FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 1441
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1264
LENGTH: 339234
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(339234)
OTHER INFORMATION: n = A,T,C or G
US-10-669-920-1264
```

```
Query Match
Best Local Similarity 50.6%; Pred. No. 45;
```

```
Matches 81; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
```

```
QY 149 GACAAATCTTAATGAACACGCTATGACGTGAAAAACAATTTGATACGAAACCTAA 208
    |||||
DB 68076 GAAACATCAATGAATGACATGAGGTATTAAGTAAAGATGATGAAATATTA 68135
    |||||
QY 209 AGGATCAACAGGAATCTGTGTGCGGTGAGATCAACGTTGCAATATTGTAAGCT 268
    |||||
DB 68136 ATCAATTCAGAGAAATGACATGTGTAGCAATCTGTTCAATTTGAATATTAATCTAGATT 68195
    |||||
QY 269 ATAAAAATGCTTCGAGTAAAGGAAAAAATTAATGA 308
    |||||
DB 68196 ATTAATTTGAAATTTTATTAAGTATCAAAATGATTAATTA 68235
    |||||
```

```
RESULT 11
US-11-266-748A-207141/C
Sequence 207141, Application US/11266748A
Publication No. US2006013463A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
```

```
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE OF INVENTION: Methods of Using the Same
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 207141
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-207141
```

```
Query Match
Best Local Similarity 55.7%; Score 33.4; DB 8; Length 1000;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 109 TATCTGTAGACGAGTGTATGCAAGCCCATAGATAGACAAATCTTAATGAACAC 168
Db 797 TAGCCATTAGTATATGTCATATCGAGCATTGAATGTGGCGAGTGAAG 738
QY 169 GGCTAGACTGCAAAACAAATTTGTATACGAAACCTTAAGCATTCAACAGAA 223
Db 737 TGTTATTAAGTGAATATATATCTGATTTTGAAACCTAGTGAATTAAGAA 683
```

```
RESULT 12
US-10-641-321-205/c
Sequence 205, Application US/10641321
GENERAL INFORMATION:
APPLICANT: Berlin, Kurt
APPLICANT: Olek, Alexander
APPLICANT: Beck, Stephan
APPLICANT: Hildmann, Thomas
APPLICANT: Lewin, Jörn
APPLICANT: Novik, Karen
TITLE OF INVENTION: Methods and Compositions for Differentiating Tissues or Cell Type
FILE REFERENCE: 47675-49
CURRENT APPLICATION NUMBER: US/10/641,321
CURRENT FILING DATE: 2003-08-12
NUMBER OF SEQ ID NOS: 221
SEQ ID NO 205
LENGTH: 4647455
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (4267840)
OTHER INFORMATION: unknown base
US-10-641-321-205
```

```
Query Match
Best Local Similarity 53.4%; Score 33.4; DB 6; Length 4647455;
```

```
Matches 70; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 176 ACTGTGAAAAACAATTTGTATGCGAAACCTTAAGCATTCAACAGAACTTGCTGTGG 235
Db 241817 ACTGTGTACAGATACCAATTCAGGACCTTATCTGAACAAATGGCTTACTGATGA 241758
QY 236 CGTGAAGTACACCGTTCGCAATATTTGTATGATTAATAAATGCTTCGAGATACGAAA 295
Db 241757 GGTGTACTGAAATATGTCCTTAAGATATGATTAATTAAGCACAATAAGAGGCAATG 241698
QY 296 AAGAAATATTT 306
Db 241697 GAAAAATTTT 241687
```

```
RESULT 13
US-11-266-748A-411280/c
Sequence 411280, Application US/11266748A
GENERAL INFORMATION:
APPLICANT: Hartin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE OF INVENTION: Methods of Using the Same
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 411280
LENGTH: 397
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (126)..(126)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120)..(120)
OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-411280
```

```
Query Match
Best Local Similarity 52.6%; Score 33.2; DB 8; Length 397;
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 329 GTCCGTAAACAGATACAGATGAGAAACGGTGGTTAAGACCTTGGAGATGTCAT 388
Db 341 GTCTTAATAACATTCAGACATGATACCAATATGATTAAGCAGATATTTTGTTCAT 282
QY 389 CGTTTCCCTTATGCTAATATCTGACGAGAAATGTTATTAAGAAAGTCTAGTATCTT 448
```

Mon Sep 18 16:01:18 2006

US-10-815-774-5.rnpbn

Db 281 CAGTTCACTAAAGTATGAGTTAAAGCCATATTATGCTTCAGTGTATCAT 222
QY 449 ACCTGTATGACAAACA 463
Db 221 GNAATTAGTTAAA 207

RESULT 14
US-11-266-748A-206397/C
; Sequence 206397, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkins, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; TITLE OF INVENTION: Method of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 206397
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-206397

Query Match 5.9%; Score 33; DB 8; Length 1000;
Best Local Similarity 50.3%; Pred. No. 3.2;
Matches 81; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 140 ATAAGATGAGACAAATCTTAATGACACGCGTATGACTGTGAAAAACAATTGTATACG 199
Db 623 ATAAGATGATGACATGGAATTAAGACACCTGAAAAACATTAATAAAAAATGTAAAGT 564
QY 200 GAAACCTAAAGGATCAACGAGAACTTGTGTGCGTGAAGCTACCACTTGCATAT 259
Db 563 CAAACCAAAACCTTAACATTAACACGCACTGACCTGACCATAGCAAGTCTATCC 504
QY 260 TGGTAAGCTATTAATAATGCTTCCGAGTAAAGGAAAAAGA 300
Db 503 TGGACACGAAAAAGTCGATTTCTTTAAATGCCAAAAAA 463

RESULT 15
US-10-953-349-36486
; Sequence 36486, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 36486
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-36486

Query Match 5.9%; Score 33; DB 6; Length 1467;
Best Local Similarity 57.9%; Pred. No. 3.9; Indels 1; Gaps 1;
Matches 77; Conservative 0; Mismatches 55

QY 372 CTCTGGGATCTGCAATCCGTTTCTCTATCTTAATCTGACGAGAAATGTTTATATACG 431
Db 59 CTACTATGATTTCTGCACTGATCTGATCTATATGAGAACTGAGCTTTACTATGACAG 118
QY 432 AAGTCTAGTAATTTTACCTGTATGA-CAACAACTGATCAATCAACTGTGTACTGA 490
Db 119 CAACCTGGAGTTGGATTCATATGATCAACGACTCAACAAATATGTCCTTGTATGA 178
QY 491 CTAAATTTGACT 503
Db 179 CCAGAGCAACT 191

Search completed: September 16, 2006, 18:38:33
Job time : 203 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2006, 14:56:20 ; Search time 3896 Seconds
(without alignments)
7965.921 Million cell updates/sec

Title: US-10-815-774-5

Perfect score: 555
Sequence: 1 ttgcgcagcagtaattccctga.....atgtatccagctgctgattt 555

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_est12:*
12: gb_est13:*
13: gb_est14:*
14: gb_est15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	41.2	7.4	706	12	CE750748
C 2	40.8	7.4	593	4	CBS584627
C 3	40.6	7.4	606	14	FR0009895
C 4	40.6	7.3	876	9	DN089815
C 5	40.4	7.3	233	1	AA767965
C 6	40	7.2	600	14	DE045970
C 7	40	7.2	825	14	DE260973
C 8	39.4	7.1	438	2	BI379434
C 9	39.4	7.1	498	2	BI379434
C 10	39.4	7.1	567	4	BM835038
C 11	39	7.0	532	2	BI893381
C 12	39	7.0	570	7	AM937387
C 13	39	7.0	570	7	AM937387
C 14	38.6	7.0	374	5	CJ349523
C 15	38.4	6.9	2553	6	AK086154
C 16	38	6.8	342	5	CJ391019
C 17	38	6.8	599	4	BM834458
C 18	38	6.8	782	5	CK861052
C 19	37.8	6.8	716	4	BM784868

C 20	37.6	6.8	729	3	BP020887
C 21	37.6	6.8	858	7	BF697486
C 22	37.6	6.8	932	14	CR240908
C 23	37.6	6.8	958	14	CR276499
C 24	37.4	6.7	401	8	CN757303
C 25	37.4	6.7	401	8	CN757304
C 26	37.4	6.7	489	3	BM885104
C 27	37.4	6.7	531	2	BM092866
C 28	37.4	6.7	899	8	CV832934
C 29	37.4	6.7	961	12	CU091135
C 30	37.2	6.7	728	14	AG164728
C 31	37.2	6.7	784	4	BM953380
C 32	37	6.7	496	12	CE347872
C 33	37	6.7	561	9	DN354639
C 34	37	6.7	689	9	DN265436
C 35	37	6.7	709	9	DN267374
C 36	37	6.7	713	9	DN267044
C 37	36.8	6.6	550	11	AQ982897
C 38	36.8	6.6	692	9	CX577073
C 39	36.8	6.6	1052	14	DU780626
C 40	36.6	6.6	555	2	BG683527
C 41	36.6	6.6	820	13	CM938378
C 42	36.6	6.6	883	7	BB898984
C 43	36.4	6.6	422	14	DR7K7S
C 44	36.4	6.6	433	12	BZ685656
C 45	36.4	6.6	491	11	AQ143297

ALIGNMENTS

RESULT 1
LOCUS CE750748/c 706 bp DNA linear GSS 30-SEP-2003
DEFINITION tigr-gss-dog-17000369607113 Dog library Canis familiaris genomic.
ACCESSION CE750748
VERSION CE750748.1 GI:37091365
KEYWORDS GSS
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 706)
AUTHORS Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Class: shotgun.
FEATURES
source location/Qualifiers
1..706
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN
Query Match 7.4%; Score 41.2; DB 12; Length 706;
Best Local Similarity 55.6%; Pred. No. 2.2;
Matches 79; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 207 AAGGATGACAGAACTTGTGCTGGAGCTACGCTTCAATATTGTAG 266
 DB 144 AGGATGACAAATGCTGTGTGATGTGAGAAAGAAATTCATGTGCACTGTGG 85
 QY 267 CTATATAAATGCTCCGAGTAACGAAAAAGATTATTGATCTTCACTATTTCAG 326
 DB 84 ATTGAAAAATTGGGACGCTACTATGAAAAATAGCATGAGATCCTCAAAATTAAAA 25
 QY 327 CGGTCTGTACAGATACCA 348
 DB 24 CAGAACTACCAAGATCCAGCA 3

RESULT 2
 CB584627
 LOCUS
 DEFINITION
 AMGNNUC:NRHY4-00187-A11-A W Rat hypothalamus (10464) Rattus
 norvegicus cDNA clone nrhy4-00187-a11 5', mRNA sequence.
 ACCESSION
 CB584627
 VERSION
 CB584627
 KEYWORDS
 EST.
 SOURCE
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Rattus.
 1 (bases 1 to 593)
 REFERENCE
 1 (bases 1 to 593)
 AUTHORS
 Amgen EST Program.
 TITLE
 Amgen Rat EST Program
 JOURNAL
 Unpublished (2003)
 COMMENT
 Contact: Dan Fitzpatrick
 Amgen, Inc
 One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 00187 row: a column: 11.
 Location/Qualifiers
 1..593
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="nrhy4-00187-a11"
 /clone_lib="W Rat hypothalamus (10464)"
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat
 hypothalamus adult female Wistar rat avg. insert size 2.3
 kb fraction 6 and 7"
 ORIGIN
 Query Match 7.4%; Score 40.8; DB 4; Length 593;
 Best Local Similarity 55.7%; Pred. No. 2.8; Indels 0; Gaps 0;
 Matches 78; Conservative 0; Mismatches 62

QY 289 ACGAAAAAAGATATTGATCTTCACTATTTCAGCGCTCTTAACAGATACCA 348
 DB 280 AATATAAACAATCAATCTTCTTCAATGTTAAAAAGTTCTGGCAAAACCAACA 339
 QY 349 TGGAGAAAGCGTTGCGTTAAACACCTCTTGGCATCTGCATCGTTTCTTTAGCTAAT 408
 DB 340 CTTAAAAAGACTGATGATTAAGCTGTCTGAATATGATGCTCTTCTTAATTGGCAAT 399
 QY 409 ACTGCAGGAAGTTTATTA 428
 DB 400 ACTTAATTAATGCTTAATTA 419

RESULT 3
 FR0009895/c
 LOCUS
 DEFINITION
 F. rubripes GSS sequence, clone 032C09AD10, genomic survey sequence.
 ACCESSION
 AL001177
 VERSION
 AL001177.1 GI:2439029
 KEYWORDS
 GSS; genome survey sequence.
 SOURCE
 Takifugu rubripes (Pugu rubripes)

ORGANISM
 Takifugu rubripes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Takifugu.
 1 (bases 1 to 606)
 REFERENCE
 1 (bases 1 to 606)
 AUTHORS
 Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrania, Y.,
 Bouchieb, N., Cottage, A., Yeo, G.S., Umrania, Y., Williams, G., and
 Brenner, S.
 TITLE
 Generation and analysis of 25 Mb of genomic DNA from the pufferfish
 Fugu rubripes by sequence scanning
 JOURNAL
 Genome Res. 9 (10), 960-971 (1999)
 PUBMED
 10523524
 COMMENT
 One pass dye-terminator sequencing of cosmid cloned genomic
 sequence.
 FEATURES
 source
 location/Qualifiers
 1..606
 /organism="Takifugu rubripes"
 /mol_type="genomic DNA"
 /db_xref="taxon:31033"
 /clone="032C09AD10"
 /clone_lib="cosmid 032C09"
 ORIGIN
 Query Match 7.4%; Score 40.8; DB 14; Length 606;
 Best Local Similarity 50.3%; Pred. No. 2.8; Indels 0; Gaps 0;
 Matches 93; Conservative 0; Mismatches 92

QY 183 AAAAAATTTGATACGAAACCTTAAGCATCAAGAACTTGTGCGCTGGAG 242
 DB 553 AGAATATTTGTAATGTGGTGNACGAAGATTCATCCCAATTTNAGCGTTTGGT 494
 QY 243 CTACACGTTGCAATATTGTAAGCTATTAATAATCTTCCGAGTAACGAAAAAGAT 302
 DB 493 AAGATATGTCGCAAAAGTGAAGCTAAATATTAATCTTAATGCGCAAGTNGAT 434
 QY 303 TATTGATCTTCACTTTTCAAGCGCTCTTAACAGATACAGATGAGAAAGCTTG 362
 DB 433 TTATCTCAATCTTATTTGAAATTTCTGTAGAGATAGATGTCATTAAGCGCGG 374
 QY 363 CGTTA 367
 DB 373 CTTGA 369

RESULT 4
 DN089815/c
 LOCUS
 DEFINITION
 JGI CABE3015.fwd NIH XGC tropoval xenopus tropicalis cDNA clone
 IMAGE:7821900 5', mRNA sequence.
 ACCESSION
 DN089815
 VERSION
 DN089815.1 GI:59756919
 KEYWORDS
 EST.
 SOURCE
 Xenopus tropicalis (western clawed frog)
 ORGANISM
 Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 876)
 REFERENCE
 1 (bases 1 to 876)
 AUTHORS
 Richardson, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,
 Brokstein, P. and Lindquist, E.A.
 TITLE
 DOE Joint Genome Institute Xenopus tropicalis EST project
 JOURNAL
 Unpublished (2004)

COMMENT

Other_ESTs: JGI CABE3015.rev
 Contact: Lindquist, E.A., Richardson, P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Robert M. Grainger
 CDNA Library Preparation: Bruce Blumberg Laboratory, University of California, Irvine
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Clone Distribution: I.M.A.G.E. Consortium/LLNL:
<http://image.llnl.gov>
 Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix 'fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Plate: CABE 0029 row: n column: 10
 High quality sequence stop: 846.
 Location/Qualifiers

FEATURES

source

```
1..876
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7821900"
/sex="Female"
/tissue_type="Ovary"
/dev_stage="Adult"
/lab_host="Electromax DH10B TI Phage Resistant cells"
/clone_1ib="NIH XCC tropoval"
/notes="Vector: PCS107; Site 1: EcoRI; Site 2: XhoI; The library was prepared from 5 ug of poly A+ RNA by oligo-dT priming
(5'-ACTAGTCGGCCGCCCTCAGCCTTCTTTTCTTTTCTTTT-3') and
Stratascript reverse transcriptase. After ligation of
EcoRI adapters (5'-AATTCGCGACGAG-3') followed by Kinaing
adapters and by XhoI digestion, the cDNA was size selected
by chromatography on Sepharose CL-2B columns and fractions
containing cDNAs larger than 1000 bp were ligated into
EcoRI/XhoI-digested PCS107. Reference for library
construction: Current Genomics 4, 635-644. Library
constructed by Michelle Tabb and Bruce Blumberg (Dept of
Developmental and Cell Biology, University of California,
Irvine)."
```

ORIGIN

Query Match 7.3%; Score 40.6; DB 9; Length 876;
 Best Local Similarity 54.3%; Pred. No. 3.5;
 Matches 82; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

```
QY 128 ATGCAAGAGCCCATAGATGAGACAAATCTTAATGAACAAGCGCTATGCTGAAAAAC 187
DB 307 ATAAAAACGCCATATGTAATAACACAGTATGATGAATACACCTTTGCTTGAAAAAT 248
QY 188 AATTTGTATAGCGAAACCTAAGGATCAACGAACTTGCTGTGAGGCTGAGCTAC 247
DB 247 ATACTTTTACTGCCATTTTAAAGTACAAACATCCCTTTGTTGTGAGAAAGGCTTATT 188
QY 248 ACGTTGCATATTTGTTAGCTATTAATAATGC 278
DB 187 AATGAGACAAATGATGAGGCTGTCTGTGTC 157
```

RESULT 5
 LOCUS AA767965 233 bp mRNA linear EST 07-FEB-1998
 DEFINITION n966a12.s1 NCI CGAP_Prl6 Homo sapiens cDNA clone IMAGE:1237630,
 mRNA sequence.
 ACCESSION AA767965
 VERSION AA767965.1 GI:2818980
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 233)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.

JOURNAL

Email: cgapsb@mail.nih.gov
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo F. Chuquí,
 M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
 Insert Length: 332 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amerham.
 Location/Qualifiers

FEATURES

source

```
1..233
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1237630"
/sex="male"
/tissue_type="tumor"
/lab_host="DH10B"
/clone_1ib="NCI CGAP_Prl6"
/notes="Organ: prostate; Vector: PAMP10; mRNA made from
invasive prostate tumor cells, cDNA made by oligo-dT
priming. Non-directionally cloned. Size-selected on
agarose gel, average insert size 600 bp. Library made by
D. Krizman, NIH."
```

ORIGIN

Query Match 7.3%; Score 40.4; DB 1; Length 233;
 Best Local Similarity 49.5%; Pred. No. 2.8;
 Matches 104; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

```
QY 18 TGATGTACCTACATTAAATCTTTTATTCATCAATAAAGAAATGCTTGGGTAACCTC 77
DB 21 TGGTTTGTAAACATACATGATATACAGATTCACATATGATCTTGTAGGCTTC 80
QY 78 TAGCGCGTCTCAGCATGATCATCTTACATATCTCTGTAAGCGATGTATGCAAGAC 137
DB 81 TTTTGTCCAAACCATGCGGTTCCATACAGAGCCTCTGACACATGTAGCTGATGAGC 140
QY 138 CCATAAGATGAGACAAATCTTAATGACAAAGCGCTATGCTGTGAAAAACAATTTGTATA 197
DB 141 AGATTAATAATGCTGATGTCATTAATTAAGGTATGCTGTAATTAATAAGCAAGTGATTT 200
QY 198 CGGAACCTTAAAGGATCAACAGAACTTG 227
DB 201 TCCAAACAGTATTCAAAAAAGATCG 230
```

RESULT 6
 LOCUS DE045970/c 600 bp DNA linear GSS 25-MAY-2005
 DEFINITION Oryzias latipes DNA, clone: olal-011004.F, genomic survey sequence.
 ACCESSION DE045970
 VERSION DE045970.1 GI:62557514
 KEYWORDS GSS.
 SOURCE Oryzias latipes (Japanese medaka)
 ORGANISM Oryzias latipes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.
 REFERENCE 1

AUTHORS Fujiyama, A., Toyoda, A., Kuroki, Y. and Sakaki, Y.
TITLE BAC end sequences of Olal Oryzias latipes library
JOURNAL Published Only in Database (2005)
REFERENCE 2 (bases 1 to 600)
AUTHORS Fujiyama, A.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2005) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suenhiro-chou, Tsunumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: atujiyam@sc.riken.jp, URL: http://stc.gsc.riken.jp/, Tel: 81-3-4212-2558, Fax: 81-3-3556-1916)
COMMENT This work was done in collaboration with Takeda, H. (1), Naruse, K. (2) and Narita, T. (3)
 (1) Department of Biological Science, University of Tokyo
 Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, JAPAN
 Phone: +81-3-5841-4431
 Fax: +81-3-5841-4993
 E-mail: htakeda.s.u-tokyo.ac.jp
 (2) Department of Biological Science, University of Tokyo
 Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, JAPAN
 Phone: +81-3-5841-4431
 Fax: +81-3-5841-4993
 E-mail: naruse.s.u-tokyo.ac.jp
 (3) Department of Biological Science, University of Tokyo
 Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, JAPAN
 Phone: +81-3-5841-4431
 Fax: +81-3-5841-4993
 E-mail: tanarita.s.u-tokyo.ac.jp
PRIMERS
Sequencing : Forward
LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 L.Site 2 : SacI
Location/Qualifiers
 1. 600
 /organism="Oryzias latipes"
 /mol_type="genomic DNA"
 /db_xref="taxon:8090"
 /clone="olal-011D04.F"
 /sex="male"
 /cell_type="whole body"
 /clone_lib="BAC end sequences of Olal Oryzias latipes library"

FEATURES
source

ORIGIN

Query Match 7.2%; Score 40; DB 14; Length 600;
 Best Local Similarity 55.9%; Pred. No. 4.7;
 Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

```

Oy 36 TTCTTATTCATCAATAAAGATCAGTCTTGCGGTACCTCTACGGCGTCTCACCATG 95
    |||||
Db 402 TTCTTTATTTAAACAAATTAACAAGCCTGCTAGTTTGGGCTGCAAGACATGCTCAACA 343
    |||||
Oy 96 CATCAATTCAGATATCTGTGAGCGAGTGTATGCAAGAGCCCATTAAGATGACAAAT 155
    |||||
Db 342 AATGCAATTCAGATTTGCAAGTTAAAGAGATTCGGCTGTTGCTATGATCAAAAA 283
    |||||
Oy 156 CTTAATGACACAGGC 171
    |||||
Db 282 ACTAAAAACAATGGC 267
    |||||
  
```

RESULT 7
 DE260973/c DE260973 825 bp DNA linear GSS 22-SEP-2005
LOCUS Oryzias latipes DNA, clone: olal-148A13.F, genomic survey sequence.
DEFINITION DE260973
ACCESSION DE260973
VERSION DE260973.1 GI:76077885

KEYWORDS GSS.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
REFERENCE 1
AUTHORS Fujiyama, A., Toyoda, A., Kuroki, Y. and Sakaki, Y.
TITLE BAC end sequences of Olal Oryzias latipes library
JOURNAL Published Only in Database (2005)
REFERENCE 2 (bases 1 to 825)
AUTHORS Fujiyama, A.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-2005) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suenhiro-chou, Tsunumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: atujiyam@sc.riken.jp, URL: http://stc.gsc.riken.jp/, Tel: 81-3-4212-2558, Fax: 81-3-3556-1916)
COMMENT This work was done in collaboration with Takeda, H. (1), Naruse, K. (2) and Narita, T. (3)
 (1) Department of Biological Science, University of Tokyo
 Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, JAPAN
 Phone: +81-3-5841-4431
 Fax: +81-3-5841-4993
 E-mail: htakeda.s.u-tokyo.ac.jp
 (2) Department of Biological Science, University of Tokyo
 Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, JAPAN
 Phone: +81-3-5841-4431
 Fax: +81-3-5841-4993
 E-mail: naruse.s.u-tokyo.ac.jp
 (3) Department of Biological Science, University of Tokyo
 Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, JAPAN
 Phone: +81-3-5841-4431
 Fax: +81-3-5841-4993
 E-mail: tanarita.s.u-tokyo.ac.jp
PRIMERS
Sequencing : Forward
LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 L.Site 2 : SacI
Location/Qualifiers
 1. 825
 /organism="Oryzias latipes"
 /mol_type="genomic DNA"
 /db_xref="taxon:8090"
 /clone="olal-148A13.F"
 /sex="male"
 /cell_type="whole body"
 /clone_lib="BAC end sequences of Olal Oryzias latipes library"

FEATURES
source

ORIGIN

Query Match 7.2%; Score 40; DB 14; Length 825;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

```

Oy 36 TTCTTATTCATCAATAAAGATCAGTCTTGCGGTACCTCTACGGCGTCTCACCATG 95
    |||||
Db 397 TTCTTTATTTAAACAAATTAACAAGCCTGCTAGTTTGGGCTGCAAGACATGCTCAACA 338
    |||||
Oy 96 CATCAATTCAGATATCTGTGAGCGAGTGTATGCAAGAGCCCATTAAGATGACAAAT 155
    |||||
Db 337 AATGCAATTCAGATTTGCAAGTTAAAGAGATTCGGCTGTTGCTATGATCAAAAA 278
    |||||
Oy 156 CTTAATGACACAGGC 171
    |||||
Db 277 ACTAAAAACAATGGC 262
    |||||
  
```


RESULT 8
BI379434/c 438 bp mRNA linear EST 26-ANG-2003
LOCUS BFLG1_001065 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG
DEFINITION or MEMGP498) Branchiostoma floridae cDNA clone MEMGP498E1916 5',
mRNA sequence.
ACCESSION BI379434
VERSION BI379434.1 GI:30914584
KEYWORDS EST.
SOURCE Branchiostoma floridae (Florida lancelet)
ORGANISM Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE 1 (bases 1 to 438)
AUTHORS Panopoulou, G., Hennig, S., Groth, D., Krause, A., Pouetka, A.J.,
Herwig, R., Vingron, M. and Lehrach, H.
TITLE New evidence for genome-wide duplications at the origin of
vertebrates using an amphioxus gene set and completed animal
genomes
JOURNAL Genome Res. 13 (6A), 1056-1066 (2003)
PUBMED 12799346
COMMENT Contact: Panopoulou G
Laboratory 145, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inestr. 63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: panopoul@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones giving the same hybridisation pattern with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well as the coordinates of the
rest of the clones assigned to the same fingerprint cluster as the
clone from which the above EST is generated is available at the
amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>
Clones and filters are distributed via the Resource Center/Primary
Database of the German Genome Project (<http://www.rzpd.de>)
PCR Primers
FORWARD: 5' CCCGAGCTTTACCTTACCTTACCTTCGGCTCG 3' (M13RSP)
BACKWARD: 5' GCTATTACGCGACGCTGGGGAAGGGGATGTC 3' (M13BSP)
Insert Length: 1200 Std Error: 0.00
Seq primer: 5'-CCGCTCCGGAATTCGCGGT-3' pSPORT3/86
High quality sequence stop: 438.
Location/Qualifiers
1..438
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="MEMGP498E1916"
/tissue_type="whole embryo"
/dev_stage="5-6 hrs (gastrula stage)"
/lab_host="E.coli, XL1 blue"
/clone_lib="Amphioxus 5-6 hrs cDNA library (Name
convention: BFLG or MEMGP498)"
/note="Vector: pSPORT1; Site 1: SalI, XbaI, EcoRI (5');
Site 2: NotI, BamHI, HindIII (3'); OligodT primed and
directionally cloned in pSPORT1 vector using a NotI
(5'-PACTAGTTCTAGATCGGAGCGGCGCGC (T)15-3' and a SalI 5'-
TCGACCCAGCGCTCG-3' adapters (Gidco BRL)."

ORIGIN

Query Match 7.1%; Score 39.4; DB 2; Length 438;
Best Local Similarity 57.9%; Pred. No. 6.5;
Matches 70; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 394 TCCCTTAATGCTAATCTGACGAGAAATGTTATTACAGAGTCTTATTTCTTACCTG 453
DB 267 TACTCTTAAGCTACACAGCAAGCAAGTAATCAAGAGCTTACAAATCTTACCTA 208

QY 454 TATGACAAATCTGATCAATACCACTGTGTACTGACTAATTTTCACGCTTCGGA 513
DB 207 TGGCATTAATTAATAAATAATGGAAGTGTCTATTAATTAATTAAGTGTGACTATA 148
QY 514 T 514
DB 147 T 147
RESULT 9
BM835568 498 bp mRNA linear EST 23-OCT-2005
LOCUS BM835568 Amphioxus Branchiostoma floridae unpublished cDNA library,
DEFINITION larva whole animal Branchiostoma floridae cDNA clone bflv055g01 3',
mRNA sequence.
ACCESSION BM835568
VERSION BM835568.1 GI:66443784
KEYWORDS EST.
SOURCE Branchiostoma floridae (Florida lancelet)
ORGANISM Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE 1 (bases 1 to 498)
AUTHORS Yu, J., Holland, L.Z., Shin, I.T., Kohara, Y., Satou, Y. and Satoh, N.
TITLE Expressed genes in Branchiostoma floridae
JOURNAL Unpublished (2005)
COMMENT Contact: Tadao Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
If you want to have a cDNA clone for this EST or if you have any
questions, please send an e-mail to Nori Satoh
(satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou
(yutaka@ascidian.zool.kyoto-u.ac.jp).
Location/Qualifiers
1..498
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="bflv055g01"
/tissue_type="larva"
/dev_stage="larva"
/clone_lib="Amphioxus Branchiostoma floridae unpublished
cDNA library, larva whole animal"

ORIGIN

Query Match 7.1%; Score 39.4; DB 4; Length 498;
Best Local Similarity 57.9%; Pred. No. 6.7;
Matches 70; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 394 TCCCTTAATGCTAATCTGACGAGAAATGTTATTACAGAGTCTTATTTCTTACCTG 453
DB 112 TACTCTTAAGCTACACAGCAAGCAAGTAATCAAGAGCTTACAAATCTTACCTA 171
QY 454 TATGACAAATCTGATCAATACCACTGTGTACTGACTAATTTTCACGCTTCGGA 513
DB 172 TGGTAAATTAATTAATAAATAATGGAAGTGTCTATTAATTAATTAAGTGTGACTATA 231
QY 514 T 514
DB 232 T 232

RESULT 10

BM835038 567 bp mRNA linear EST 22-OCT-2005
LOCUS BM835038 Amphioxus Branchiostoma floridae unpublished cDNA library,
DEFINITION larva whole animal Branchiostoma floridae cDNA clone bflv053o19 3',
mRNA sequence.

ACCESSION	BW635038
VERSION	BW635038.1
GI:	66443254
KEYWORDS	EST.
SOURCE	Branchiostoma floridae (Florida lancelet)
ORGANISM	Branchiostoma floridae Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.
REFERENCE	1 (bases 1 to 567) Yu, J., Holland, L.Z., Shin-I, T., Kohara, Y., Satou, Y. and Satoh, N. Expressed genes in Branchiostoma floridae Unpublished (2005)
AUTHORS	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 111 Yata, Mishima, Shizuoka 411-8540, Japan Tel.: 81-559-81-6856 Fax: 81-559-81-6855
JOURNAL	Email: tshin@genes.nig.ac.jp If you want to have a cDNA clone for this EST or if you have any questions, please send an e-mail to Nori Satoh (satoha@cridian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou (yutake@cridian.zool.kyoto-u.ac.jp).
COMMENT	Location/Qualifiers
FEATURES	1..567
source	/organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="taxon:7739" /clone="bf1v053o19" /tissue.type="whole animal" /dev_stage="larva" /clone_id="Amphioxus Branchiostoma floridae unpublished cDNA library, larva whole animal"
ORIGIN	
Query Match	7.1%; Score 39.4; DB 4; Length 567;
Best Local Similarity	57.9%; Pred. No. 6.9;
Matches	70; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
Dn	394 TCCTCTTATGCTAATAGTCGCGAGGAATGTTATTACAGAAGTCTAGTAATCTTACTGTG 453 114 TACTCTTAGCTACCAACAAGAACAGTAATAATCAAGCCTCTTACAAATATPACCTA 173
Dp	454 TATGCACAAACATCATCAATATCCAACTGTGACTACGACTAAATTTTCACGCTTTCCGGA 513 174 TGGTATPATATPATAATTAATTTGGAGTGTTGCTCTATATATATTAAGTGTATCTATA 233
Dn	514 T 514 234 T 234
RESULT 11	
LOCUS	BI693381
DEFINITION	sa165dl1.y1 Gm-cl068 glycine max cDNA clone GENOME SYSTEMS CLONE
ACCESSION	BI693381
VERSION	BI693381.1
KEYWORDS	EST.
SOURCE	Glycine max (soybean)
ORGANISM	Glycine max Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; SpERMatoPhyta; Magnoliophyta; euDicoTYledons; core euDicoTYledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE	1 (bases 1 to 532) Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McGann, R., Waterston, R. and Wilson, R.
TITLE	Public Soybean EST Project

FEATURES	source
<p>Unpublished (1999) Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watscn.wustl.edu</p> <p>When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4153; email: info@biogeneticservices.com) Seq primer: -40RP from Gldco High quality sequence stop: 425. Location/Qualifiers</p>	<p>1..532</p> <p>/organism="Glycine max"</p> <p>/mol type="mRNA"</p> <p>/cultivar="Williams 82"</p> <p>/db xref="taxon:3847"</p> <p>/clone="GENOME SYSTEMS CLONE ID: Gm-cl068-3621"</p> <p>/tissue type="Leaf, drought stressed, 1 month old plants, greenhouse grown"</p> <p>/lab host="DH10B"</p> <p>/clone_lib="Gm-cl068"</p> <p>/note="Vector: pBluescript II SK-, Site_1: EcoRI, Site_2: XhoI; The cDNA library was constructed from mRNA isolated from drought stressed leaf tissue of the cultivar Williams 82. The month old greenhouse grown plants were deprived of water for 3 days prior to harvesting the stressed leaf tissue. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."</p>
ORIGIN	
<p>Query Match 7.0%; Score 39; DB 2; Length 532; Best Local Similarity 68.4%; Pred. No. 8.9; Matches 54; Conservative 0; Mismatches 25; Indels 0; Gaps 0;</p>	
OY	34 AATCTTTATTCATCAATPAAGAATCGCTTGGCTACCTTAGCGCTCTCAACA 93
DB	5 AATCTTTTGGTCCAAACAAACGATGATGGTGGCTTCTTCTACGAGATGCTCAACA 64
OY	94 TGCATCACATTCAGATATTC 112
DB	65 TGATTCATCTTCACATTAC 83
RESULT 12	
LOCUS	AM937387/c
DEFINITION	QVA-DT0043-180100-061-c07 DT0043 Homo sapiens cDNA, mRNA sequence.
ACCESSION	AM937387
VERSION	AM937387.1 GI:8112810
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	<p>1 (bases 1 to 570) 570 bp mRNA linear EST 30-MAY-2000</p> <p>Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.</p>

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=kt2=QV3-DT0043-180
100-061-c07kt3=2000-01-18kt4=1)
Seq primer: puc 18 forward
High quality sequence stop: 569.

FEATURES
source
1..570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="DT0043"
/note="Organ: denis drash; Vector: puc18; Site 1: Smal; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 7.0%; Score 39; DB 7; Length 570;
Best Local Similarity 52.8%; Pred. No. 9;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

182 AAAACAATTGTATACGGAACCTTAAGCATCAACAGAACTGCTGTGGCGTGA 241
|||||
480 AAAAGAGTGTATATAGGAATACTGTACTTGTGATTTAGAACTTATGATTTAA 421
242 GCTACACGTCGATATTTGTAAGCTATATAAATCTCCGAGTACGAAAAAGAA 301
|||||
420 GCTACCTGTATATAGGGGTAGGCTAGAGTACCAAGCAAGAAAAAGGTAATTGG 361
302 TTATTGATCCTTCATATTTTCAAGCGGCTCTGTAAAG 340
|||||
360 CTAAGAAATTCACAAATTTAAGTGTCTTTAAAAAG 322

RESULT 13
AM937408/c 570 bp mRNA linear EST 30-MAY-2000
LOCUS QV3-DT0043-190100-061-c07 DT0043 Homo sapiens cDNA, mRNA sequence.
DEFINITION AM937408
ACCESSION AM937408.1 GI:8112831
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo
1 (bases 1 to 570)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=kt2=QV3-DT0043-190
100-061-c07kt3=2000-01-19kt4=1)
Seq primer: puc 18 forward
High quality sequence stop: 569.

FEATURES
source
1..570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="DT0043"
/note="Organ: denis drash; Vector: puc18; Site 1: Smal; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 7.0%; Score 39; DB 7; Length 570;
Best Local Similarity 52.8%; Pred. No. 9;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

182 AAAACAATTGTATACGGAACCTTAAGCATCAACAGAACTGCTGTGGCGTGA 241
|||||
480 AAAAGAGTGTATATAGGAATACTGTACTTGTGATTTAGAACTTATGATTTAA 421
242 GCTACACGTCGATATTTGTAAGCTATATAAATCTCCGAGTACGAAAAAGAA 301
|||||
420 GCTACCTGTATATAGGGGTAGGCTAGAGTACCAAGCAAGAAAAAGGTAATTGG 361
302 TTATTGATCCTTCATATTTTCAAGCGGCTCTGTAAAG 340
|||||
360 CTAAGAAATTCACAAATTTAAGTGTCTTTAAAAAG 322

RESULT 14
CJ349523/c 374 bp mRNA linear EST 15-JUN-2005
LOCUS CJ349523 Molgula tectiformis unpublished cDNA library, cleaving embryo Molgula tectiformis cDNA clone mcl013K20 5', mRNA sequence.
DEFINITION CJ349523
ACCESSION CJ349523
VERSION CJ349523.1 GI:67789671
KEYWORDS EST.
SOURCE Molgula tectiformis
ORGANISM Molgula tectiformis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Stolidobranchia; Molgulidae; Molgula.
1 (bases 1 to 374)
Gyoja, F., Satou, Y. and Satoh, N.
Expressed genes in Molgula tectiformis
Unpublished (2005)
Contact: Tadao Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
When you want to obtain this EST clone, please send an e-mail to Nori Satoh (satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to yutaka

FEATURES
source
Satou (yutaka@ascidian.zool.kyoto-u.ac.jp).
Location/Qualifiers
1..374
/organism="Molgula tectiformis"
/mol_type="mRNA"
/db_xref="taxon:30286"
/clone="mtc1013k20"
/tissue_type="whole animal"
/dev_stage="cleaving embryo"
/clone_lib="Molgula tectiformis unpublished cDNA library,
cleaving embryo"

ORIGIN
Query Match 7.0%; Score 38.6; DB 5; Length 374;
Best Local Similarity 50.8%; Pred. No. 11;
Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 272 AAAATGCTTCGGAGTACGGAAGAAAGAAATTTGATCTCTTCACTATTTCAGAGCGTC 331
DB 352 AGAGTCTTTTCCAAAGTAAAGTTAAATGAATCCATTTGCTTCACGAAATGGAATCTTTTC 293
QY 332 CTGTAAAGATACAGCATGAGAAAGCGCTTGCGTTACACCTTCGCGATCTGACCTCG 391
DB 292 TGTTAAGATGAAATTAATGATCATATATTTGAAGTAAACAGTACGAGGACGACCTT 233
QY 392 TTTCCTCTTATGCTAATATCTGACGAAATGTTTATTAACAGAGTCTTATTTTAC 451
DB 232 TGTATATGGGCTTTTACACAGCTGCTTCTCAATGATGATTAATCTTTTAAATGTTTCT 173
QY 452 T 452
DB 172 T 172

RESULT 15
AK086154 2553 bp mRNA linear HTC 02-SEP-2005
LOCUS Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched
DEFINITION library, clone:D930008124 product:unclassifiable, full insert
sequence.
ACCESSION AK086154 GI:26103259
VERSION AK086154.1 GI:26103259
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
PUBMED Genome Res. 10 (10), 1617-1630 (2000)
REFERENCE 11042159

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashinagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, N., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format
JOURNAL sequencing pipeline with 384 multichannel sequencer
PUBMED Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE 11076861

AUTHORS
TITLE The RIKEN Genome Exploration Research Group Phase II Team and the
JOURNAL PANTOM Consortium.
REFERENCE Functional annotation of a full-length mouse cDNA collection
AUTHORS Nature 409, 685-690 (2001)

TITLE
JOURNAL The PANTOM Consortium, the RIKEN Genome Exploration Research Group
REFERENCE Phase I and II Team.
AUTHORS Analysis of the mouse transcriptome based on functional annotation
OF 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL RIKEN Genome Exploration Research Group, Genome Science Group
REFERENCE (Genome Network Core Team) and the PANTOM Consortium.
AUTHORS Antisense Transcription in the Mammalian Transcriptome
Science 309, 1564-1566 (2005)

TITLE
JOURNAL The PANTOM Consortium, Riken Genome Exploration Research Group and
REFERENCE The Transcriptional Landscape of the Mammalian Genome
AUTHORS Science 309, 1559-1563 (2005)

TITLE
JOURNAL The PANTOM Consortium, Riken Genome Exploration Research Group and
REFERENCE The Transcriptional Landscape of the Mammalian Genome
AUTHORS Science 309, 1559-1563 (2005)

TITLE
JOURNAL Submitted (16-Apr-2002) Yoshihide Hayashizaki, The Institute of
REFERENCE Physical and Chemical Research (RIKEN), Laboratory for Genome
AUTHORS Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:htp://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
Direct Submission

COMMENT
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:htp://genome.gsc.riken.jp/
URL:htp://pantom.gsc.riken.jp/
Location/Qualifiers

FEATURES
source
1..2553
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D930008124"
/db_xref="taxon:10090"
/clone="D930008124"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="15 days embryo"
1..2553
/note="putative
unclassifiable"

ORIGIN
Query Match 6.9%; Score 38.4; DB 6; Length 2553;
Best Local Similarity 57.5%; Pred. No. 20;
Matches 69; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 410 CTGACGAATGTTTATTAACAGAGCTCTAGTATCTTACCGTGTGACCAATCTG 469
DB 714 CTGTAGGAATTTCTGTACAGTCTGTATGATCTTTTCAATATGATGATCTG 773

Qy 470 TCAATACCACTGTGACTGACTAAATTTTCACTGCTTCCGGATGTTCTCTTCACTG 529
Db 774 AAATAAGATATTGTGCTTCTTCCCACTTCTATCCCTGTGTCTCTTCACTG 833.

Search completed: September 16, 2006, 17:03:14
Job time : 3903 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 06:14:30 ; Search time 197 Seconds
(without alignments)
429.366 Million cell updates/sec

Title: US-10-815-774-6

Perfect score: 991
Sequence: 1 LASVIPDVATINSLEFNQIKN.....FSLSGCSPSPAPDVSSCGF 185

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq 8:*

- 1: geneseqp1980s:*
- 2: geneseqp1980s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*
- 10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	185	4 AAB84386	Aab84386 Amino aci
2	991	100.0	320	4 AAB84387	Aab84387 Amino aci
3	991	100.0	320	9 AEE01516	Aee01516 Corynefor
4	874	88.2	185	3 AAY44582	Aay44582 Chryseoba
5	874	88.2	319	3 AAY44583	Aay44583 Chryseoba
6	125.5	12.7	274	9 AEB40953	Aeb40953 L. pneumo
7	125.5	12.7	286	9 AEB37640	Aeb37640 L. pneumo
8	108	10.9	20	4 AAB84385	Aab84385 Internal
9	96	9.7	20	4 AAB84384	Aab84384 N-termina
10	92	9.3	571	7 ABO69131	AbO69131 Pseudomon
11	91	9.2	594	3 AAG52479	Aag52479 Arabidops
12	91	9.2	594	3 AAG52451	Aag52451 Arabidops
13	91	9.2	594	3 AAB91568	Abb91568 Herbicida
14	90.5	9.1	292	8 ADY09066	Ady09066 Plant ful
15	89.5	9.0	927	8 ADN23916	Adn23916 Bacterial
16	89.5	9.0	927	8 ADN23918	Adn23918 Bacterial
17	86	8.7	1219	7 AAE37900	Aae37900 Rice homo
18	85.5	8.6	442	8 ADX72461	Adx72461 Plant ful
19	85.5	8.6	477	4 AAB30815	Aab30815 Amino aci
20	85.5	8.6	477	10 AEE28467	Aee28467 Yeast unc
21	85.5	8.6	477	10 AEE72832	Aee72832 S. cerevi
22	85.5	8.6	961	8 ADY23617	Ady23617 Plant ful
23	85.5	8.6	1142	8 ADY06625	Ady06625 Plant ful

ALIGNMENTS

24	85	8.6	264	3 AAG25396	Aag25396 Arabidops
25	85	8.6	292	3 AAG25395	Aag25395 Arabidops
26	85	8.6	294	3 AAG25394	Aag25394 Arabidops
27	85	8.6	594	10 AEF80099	Aef80099 A. thalia
28	85	8.6	907	10 AEE60297	Aee60297 Cat chlam
29	85	8.6	907	10 AEE60508	Aee60508 Cat chlam
30	84	8.5	3726	4 ABB63947	Abb63947 Drosophila
31	83.5	8.4	459	9 ABB93408	Abb93408 M. xanthu
32	82	8.3	703	8 ADA44005	Ada44005 Bacterial
33	82	8.3	927	2 AAY35054	Aay35054 Chlamydia
34	82	8.3	928	2 AAM88423	Aam88423 Chlamydia
35	82	8.3	928	2 AAY90239	Aay90239 Chlamydia
36	82	8.3	928	5 ABB90542	Abb90542 Chlamydia
37	82	8.3	928	9 ADW65028	Adw65028 C. pneumo
38	82	8.3	929	9 AEC95710	Aec95710 C. pneumo
39	81.5	8.2	2360	6 ABB56972	Abb56972 B. maxima
40	81	8.2	270	3 AAG21040	Aag21040 Arabidops
41	81	8.2	304	3 AAG21039	Aag21039 Arabidops
42	81	8.2	346	4 ABB71526	Abb71526 Drosophila
43	81	8.2	486	3 AAG21038	Aag21038 Arabidops
44	81	8.2	623	8 ADP04503	Adp04503 Sea squir
45	81	8.2	1137	2 AAR42081	Aar42081 Impatiens

RESULT 1
ID AAB84386 standard; protein; 185 AA.
XX
AC AAB84386;
XX
DT 06-AUG-2003 (revised)
DT 22-AUG-2001 (first entry)
XX
DE Amino acid sequence of a protein-deamidating enzyme.
XX
KW Protein-deamidating enzyme; mineral absorption; food allergy; dough;
KW bakery; confectionery.
XX
OS Chryseobacterium sp.
XX
PN EP1106696-A1.
XX
PD 13-JUN-2001.
XX
PF 04-DEC-2000; 2000EP-00310768.
XX
PR 03-DEC-1999; 99JP-00345044.
XX
PA (AMANO-) AMANO ENZYME INC.
XX
PI Yamaguchi S;
XX
DR WPI: 2001-376907/40.
DR N-PSDB; AAF90280.
XX
PT New enzyme for use in e.g. bakery has an ability to deamidate amido
PT groups in a protein.
XX
PS Claim 4; Page 22; 43pp; English.
XX
CC The present sequence represents a protein-deamidating enzyme from
CC Cryeobacterium sp. number 9670. The enzyme is able to deamidate amido
CC groups in a protein by directly acting upon the amido groups without
CC cutting peptide bonds and without cross-linking the protein. The enzyme
CC thus reduces the mineral sensitivity of the protein and increases the
CC soluble mineral content in the protein-mineral solution, improving the
CC absorption of minerals in the human body. The enzyme reduces or removes
CC the toxicity and allergenic property of the protein in a food e.g.
CC allergy. The enzyme is useful for the improvement of dough in the field
CC of bakery and confectionery e.g. for the production of crackers,

CC biscuits, cookies, pizza pies or crusts of pie; in producing soybean
 CC protein products, in various food articles e.g. meat or fish products and
 CC noodles; and for improving functionality of plant or animal protein.
 CC (updated on 06-AUG-2003 to correct OS field.)

XX Sequence 185 AA;

Query Match 100.0%; Score 991; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 6,7e-94;
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASVTPDVATINSLFNOIKNOSCGTSTASPCITFRYPVDGCVARAHKROILMNGYDC 60
 DB 1 LASVTPDVATINSLFNOIKNOSCGTSTASPCITFRYPVDGCVARAHKROILMNGYDC 60
 QY 61 EKQFYGNLKAATGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSSGPTDTAMRNA 120
 DB 61 EKQFYGNLKAATGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSSGPTDTAMRNA 120
 QY 121 CVNTSCGSASVSSYANTAGNVYRSPNSYLYDNNLINTNCVLTFRSLSGCSPSPADPV 180
 DB 121 CVNTSCGSASVSSYANTAGNVYRSPNSYLYDNNLINTNCVLTFRSLSGCSPSPADPV 180
 QY 181 SSCGF 185
 DB 181 SSCGF 185

RESULT 2
 AAB84387
 ID AAB84387 standard; protein; 320 AA.

XX AAB84387;
 XX 06-AUG-2003 (revised)
 DT 22-AUG-2001 (first entry)

XX Amino acid sequence of a protein-deamidating enzyme.

XX Protein-deamidating enzyme; mineral absorption; food allergy; dough;
 KM bakery; confectionery.

OS Chryseobacterium sp.

XX Key Location/Qualifiers
 FT Region 1..135
 XX /note= "prepro region"

PN EPI106696-A1.

XX 13-JUN-2001.

XX 04-DEC-2000; 2000EP-00310768.

XX 03-DEC-1999; 99JP-00345044.

XX (AMANO-) AMANO ENZYME INC.

XX Yamaguchi S;

XX WPI; 2001-376907/40.

DR N-PSDB; AAF90281.

XX New enzyme for use in e.g. bakery has an ability to deamidate amido
 PT groups in a protein.
 XX Example 11; Page 24; 43pp; English.

XX The present sequence represents a protein-deamidating enzyme from
 CC Cryseobacterium sp. number 9670. The enzyme is able to deamidate amido
 CC groups in a protein by directly acting upon the amido groups without
 CC cutting peptide bonds and without cross-linking the protein. The enzyme
 CC thus reduces the mineral sensitivity of the protein and increases the

CC soluble mineral content in the protein-mineral solution, improving the
 CC absorption of minerals in the human body. The enzyme reduces or removes
 CC the toxicity and allergenic property of the protein in a food e.g.
 CC allergy. The enzyme is useful for the improvement of dough in the field
 CC of bakery and confectionery e.g. for the production of crackers,
 CC biscuits, cookies, pizza pies or crusts of pie; in producing soybean
 CC protein products, in various food articles e.g. meat or fish products and
 CC noodles; and for improving functionality of plant or animal protein.
 CC (updated on 06-AUG-2003 to correct OS field.)

XX Sequence 320 AA;

Query Match 100.0%; Score 991; DB 4; Length 320;
 Best Local Similarity 100.0%; Pred. No. 1.4e-93;
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASVTPDVATINSLFNOIKNOSCGTSTASPCITFRYPVDGCVARAHKROILMNGYDC 60
 DB 136 LASVTPDVATINSLFNOIKNOSCGTSTASPCITFRYPVDGCVARAHKROILMNGYDC 195
 QY 61 EKQFYGNLKAATGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSSGPTDTAMRNA 120
 DB 196 EKQFYGNLKAATGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSSGPTDTAMRNA 255
 QY 121 CVNTSCGSASVSSYANTAGNVYRSPNSYLYDNNLINTNCVLTFRSLSGCSPSPADPV 180
 DB 256 CVNTSCGSASVSSYANTAGNVYRSPNSYLYDNNLINTNCVLTFRSLSGCSPSPADPV 315
 QY 181 SSCGF 185
 DB 316 SSCGF 320

RESULT 3
 AEE01516
 ID AEE01516 standard; protein; 320 AA.

XX AEE01516;

XX 26-JUN-2006 (first entry)

XX Coryneform protein production system-related protein, SEQ ID NO:4.

XX protein production; cell culture; genetically engineered microorganism.

XX Unidentified.

XX WO2005103278-A1.

XX 03-NOV-2005.

XX 20-APR-2005; 2005MO-JP007518.

XX 20-APR-2004; 2004JP-00124196.

XX 13-JAN-2005; 2005JP-00005896.

XX (AJIN) AJINOMOTO CO INC.

XX Date M, Kikuchi Y, Itaya H, Nakamura N;

DR WPI; 2005-748042/76.

DR N-PSDB; AEE01515.

XX Producing heterologous protein by culturing Coryneform bacterium having
 PT expression gene construct comprising promoter sequence, nucleic acid
 PT encoding Tat system-dependent signal peptide region and nucleic acid
 PT encoding foreign protein.

XX Disclosure; SEQ ID NO 4; 47pp; Japanese.

XX The new invention relates to a method of producing heterologous protein,
 CC by culturing a Coryneform bacterium having an expression gene construct
 CC carrying, in the direction from the 5'-terminal to 3'-terminal, a

CC promoter sequence functioning in the bacterium, a nucleic acid sequence
 CC encoding a Tat system-dependent signal peptide region and a nucleic acid
 CC sequence encoding a foreign protein, and allowing the bacterium to
 CC produce and secrete the protein. The signal peptide of (M1) comprises SEQ
 CC ID NO: 28-32. The signal peptide is a signal peptide of isomaltolactan,
 CC which comprises SEQ ID NO: 6, or of trimethylamine N-oxidoreductase,
 CC which comprises SEQ ID NO: 8. The bacteria comprises gene that encodes
 CC Tat system secretion peptide such as tatA, tatB, tatC or tatE. (M1) is
 CC useful for producing heterologous protein e.g. glucanase or
 CC isomaltodextran. (M1) enables efficient production of heterologous
 CC protein. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from Wipo at ftp.wipo.int/pub/published-pat sequences. The present
 CC sequence is a protein related to the invention.

XX SQ Sequence 320 AA;

Query Match 100.0%; Score 991; DB 9; Length 320;
 Best Local Similarity 100.0%; Pred. No. 1.4e-93;
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASVIPDVATLNSLFNQIKQSCGTSTASSPCITFRYPVDCYARAHKMQILMANGYDC 60
 DB 136 LASVIPDVATLNSLFNQIKQSCGTSTASSPCITFRYPVDCYARAHKMQILMANGYDC 195
 QY 61 EKQFVYGNLKASTGTCCVAMSYHVALIVSYKASGVTEKRIIDPSLFSSGPTVTARMA 120
 DB 196 EKQFVYGNLKASTGTCCVAMSYHVALIVSYKASGVTEKRIIDPSLFSSGPTVTARMA 255
 QY 121 CVNTSCGASVSSYANTAGNVYRSPNSLYLDNNLINTNCVLTKEFLSGCSPAPDV 180
 DB 256 CVNTSCGASVSSYANTAGNVYRSPNSLYLDNNLINTNCVLTKEFLSGCSPAPDV 315
 QY 181 SSCGF 185
 DB 316 SSCGF 320

RESULT 4

AAV44582
 ID AAV44582 standard; protein; 185 AA.

AC AAV44582;
 XX
 DT 12-SEP-2003 (revised)
 DT 04-APR-2000 (first entry)

DE Chryseobacterium gleum protein-deamidating enzyme.

XX Protein deamidating enzyme; soil bacterium; deamidation activity;
 KW amido group; carboxyl group; ammonia; cross linking; peptide bond;
 KW protein engineering; surface hydrophobicity; toxicity; allergic;
 KW mineral sensitivity; calcium; absorption; mineral enhancing agent.

OS Chryseobacterium gleum; 'JCM 2410'.

XX EP976829-A2.

XX 02-FEB-2000.

XX 04-JUN-1999; 99EP-00304367.

XX 04-JUN-1998; 98JP-00173940.

XX (AMANO) AMANO PHARM KK.

XX Yamaguchi S, Matsuura A;

XX MPI: 2000-11852/11.

XX N-PSDB; AA249494.

PT New enzyme for modifying and improving the function of proteins and/or
 PT peptides has deamidating activity without causing cross linking.

XX PS Claim 5; Page 24; 57pp; English.

XX The present amino acid sequence is the protein-deamidating enzyme,
 CC isolated from a new strain of soil bacterium, Chryseobacterium gleum JCM
 CC 2410. This enzyme exerts the deamidation activity by directly acting upon
 CC side chain amido groups in the protein in bonded state and releasing side
 CC chain carboxyl groups and ammonia. It can deaminate high molecular weight
 CC proteins, without cross linking and cleavage of peptide bonds, to improve
 CC protein function. This sequence is used for protein engineering, to cause
 CC an increase in surface hydrophobicity and improve the function of a plant
 CC or animal protein. It can also be used to remove or reduce toxicity or
 CC allergic property of proteins in food, decrease mineral sensitivity of
 CC protein, to allow greater absorption into the body and to solubilise
 CC calcium for use in drinks and mineral enhancing agents. (Updated on 12-
 CC SEP-2003 to standardise OS field)

XX SQ Sequence 185 AA;

Query Match 88.2%; Score 874; DB 3; Length 185;
 Best Local Similarity 87.4%; Pred. No. 8.3e-82;
 Matches 160; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 3 SVIPDVATLNSLFNQIKQSCGTSTASSPCITFRYPVDCYARAHKMQILMANGYDC 62
 DB 3 SVIPDVATLNSLFNQIKQSCGTSTASSPCITFRYPVDCYARAHKMQILMANGYDC 62
 QY 63 QFVYGNLKASTGTCCVAMSYHVALIVSYKASGVTEKRIIDPSLFSSGPTVTARMA 122
 DB 63 QFVYGNLKASTGTCCVAMSYHVALIVSYKASGVTEKRIIDPSLFSSGPTVTARMA 122
 QY 123 NTSCGASVSSYANTAGNVYRSPNSLYLDNNLINTNCVLTKEFLSGCSPAPDV 182
 DB 123 NTSCGASVSSYANTAGNVYRSPNSLYLDNNLINTNCVLTKEFLSGCSPAPDV 182
 QY 183 CGF 185
 DB 183 CGF 185

RESULT 5

AAV44583
 ID AAV44583 standard; protein; 319 AA.

AC AAV44583;
 XX

DT 12-SEP-2003 (revised)
 DT 04-APR-2000 (first entry)

DE Chryseobacterium gleum protein-deamidating prepro-enzyme.

XX Protein deamidating prepro-enzyme; soil bacterium; deamidation activity;
 KW amido group; carboxyl group; ammonia; cross linking; peptide bond;
 KW protein engineering; surface hydrophobicity; toxicity; allergic;
 KW mineral sensitivity; calcium; absorption; mineral enhancing agent.

OS Chryseobacterium gleum; 'JCM 2410'.

XX EP976829-A2.

XX Key Location/Qualifiers

XX 1.134
 FT /label= Prepro region
 FT /note= "Comprises the Pre and Pro region"

XX 1.21
 FT /label= Signal peptide
 FT /note= "Corresponds to the Pre region"

XX 22.134
 FT /note= "Corresponds to the Pro region"

XX 135.319
 FT /label= Mature protein deamidating enzyme
 FT /note= "Deamidates amido groups in a protein"

XX EP976829-A2.

PD 02-FEB-2000.
 XX 04-JUN-1999; 99EP-00304367.
 XX 04-JUN-1998; 98JP-00173940.
 XX (AMANO) AMANO PHARM KK.
 XX Yamaguchi S, Matsuura A;
 PI WPI; 2000-118552/11.
 DR N-PSDB; AA249495.
 XX
 PT New enzyme for modifying and improving the function of proteins and/or
 PT peptides has deamidating activity without causing cross linking.
 PS Example 26; Page 26; 57pp; English.
 XX
 CC The present amino acid sequence is the protein-deamidating prepro-enzyme,
 CC isolated from a new strain of soil bacterium, Chryseobacterium gleum JCM
 CC 2410. This enzyme exerts the deamidation activity by directly acting upon
 CC side chain amino groups in the protein in bonded state and releasing side
 CC chain carboxyl groups and ammonia. It can deaminate high molecular weight
 CC proteins, without cross linking and cleavage of peptide bonds, to improve
 CC protein function. This sequence is used for protein engineering, to cause
 CC an increase in surface hydrophobicity and improve the function of a plant
 CC or animal protein. It can also be used to remove or reduce toxicity of
 CC allergenic property of proteins in food, decrease mineral sensitivity of
 CC protein, to allow greater absorption into the body and to solubilise
 CC calcium for use in drinks and mineral enhancing agents. (Updated on 12-
 CC SEP-2003 to standardise OS field)
 XX
 SQ Sequence 319 AA;

Query Match 88.2%; Score 874; DB 3; Length 319;
 Best Local Similarity 87.4%; Pred. No. 1.7e-81;
 Matches 160; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

```

QY 3 SVIPDVATLNSLFNQIKNOSCGTSTASSPCTFRFPVPGCYARAKRKQILMANGYDCEK 62
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 137 SVIPDLATLNSLFQIKNGACGTSTASSPCTFRFPVPGCYARAKRKQILMANGYDCEK 196

QY 63 QFVYGNLKAATGTCVAVSHVAVILVSYKASGVTERRIIDPSLFSSGPTDPAWRVACV 122
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 197 QFVYGNLRASYGTCCVSVVHVAVILVSKASGIVERRIIDPSLFSSGPTDPAWRVACT 236

QY 123 NTSGSASVSSYANTAGNVYRSPNSLYDNNLINTNCVLTKEFSLSGSPSPAPDVSS 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 257 NTSGSASVSSYANTAGNVYRSPNSLYDNNLINTNCVLTKEFSLSGSPSPAPDVSS 316

QY 183 CGF 185
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 317 CGF 319
```

RESULT 6
 AEB40953
 ID AEB40953 standard; protein; 274 AA.

XX AEB40953;
 XX AC
 XX 08-SEP-2005 (first entry)
 XX DE L. pneumophila protein SEQ ID NO 5285.
 XX KM detection; infection; Antibacterial; Vaccine.
 XX OS Legionella pneumophila.
 XX PN WO2005049642-A2.
 XX PD 02-JUN-2005.
 XX

PF 23-SEP-2004; 2004MO-IB003578.
 XX 21-NOV-2003; 2003FR-00013687.
 XX (INSP) INST PASTEUR.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX (UTLY-) UNIV LYON 1 BERNARD CLAUDE.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX Buchrieser C, Tichit M, Etienne J, Ma L, Cazalec C, Glaeser P;
 PI Rusniok C, Bouchier C, Zidane N, Magnier A, Kunet F, Vandenesch F;
 PI Jarraud S;
 XX WPI; 2005-388305/40.
 DR
 XX
 PT New genome of Legionella pneumophila Paris strain and derived
 PT polypeptides, useful for detection or identification of the strain and
 PT for treatment and prevention of infections.
 PS Claim 3; SEQ ID NO 5285; 660pp; English.

XX The invention relates to an isolated or purified nucleotide sequences (I)
 CC from Legionella pneumophila Paris strain. (I), and their related
 CC sequences or fragments, are useful as primers and probes for detection
 CC and amplification, including differentiation between the Paris and
 CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
 CC (hybrid) polypeptides (II). (II) are also useful for preparation of
 CC specific antibodies (Ab), also used for detection/identification of
 CC Legionella, and some (I), specifically those involved in synthesis of
 CC surface proteins, are targets for identification of inhibitors. (II), or
 CC vectors that contain (I), are useful as vaccines and immunogenic
 CC compositions, for treatment and prevention of infections by L.
 CC pneumophila. The present sequence represents the amino acid sequence of a
 CC L. pneumophila protein.
 XX
 SQ Sequence 274 AA;

Query Match 12.7%; Score 125.5; DB 9; Length 274;
 Best Local Similarity 31.2%; Pred. No. 0.00035;
 Matches 43; Conservative 22; Mismatches 36; Indels 37; Gaps 8;

```

QY 3 SVIPDVAT--LNSLFNQIKNOSCGTSTASSP-----CTFRFPVPGCYARAKRKQI 52
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 72 STVPKVSYSDELTMNMLIRD---TRFLYSPDKDPFORRISWLYPDGCFRALSLRIK 127

QY 53 LMANGVDC-----KQFYGNLK-----ASTGTCVAVSHVAVILVSYKASGVTERRI 101
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 128 L-----DSHFVYIPAKIFAFGDEMOTPYSSBGT--VSMWYHVASVAVNYMGTI-----YV 175

QY 102 IDPSLFSSGPTDPAWRN 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 176 LDPAAKPEBGMULDWYN 193
```

RESULT 7
 AEB37640
 ID AEB37640 standard; protein; 286 AA.

XX AEB37640;
 XX AC
 XX 08-SEP-2005 (first entry)
 XX DE L. pneumophila protein SEQ ID NO 1972.
 XX KM detection; infection; Antibacterial; Vaccine.
 XX OS Legionella pneumophila.
 XX PN WO2005049642-A2.
 XX PD 02-JUN-2005.
 XX PF 23-SEP-2004; 2004MO-IB003578.

XX Example 11; Page 20; 43pp; English.

XX The present sequence represents a peptide fragment of a protein-
CC deamidating enzyme from *Cryoseobacterium* sp. number 9670. The enzyme is
CC able to deamidate amide groups in a protein by directly acting upon the
CC amide groups without cutting peptide bonds and without cross-linking the
CC protein. The enzyme thus reduces the mineral sensitivity of the protein
CC and increases the soluble mineral content in the protein-mineral
CC solution, improving the absorption of minerals in the human body. The
CC enzyme reduces or removes the toxicity and allergenic property of the
CC protein in a food e.g. allergy. The enzyme is useful for the improvement
CC of dough in the field of bakery and confectionery e.g. for the production
CC of crackers, biscuits, cookies, pizza pies or crusts of pie; in producing
CC soybean protein products, in various food articles e.g. meat or fish
CC products and noodles; and for improving functionality of plant or animal
CC protein. (Updated on 06-AUG-2003 to correct OS field.)

XX
SQ Sequence 20 AA;

Query Match 9.7%; Score 96; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASVIPVATLNSLENOIKN 20
Db 1 LASVIPVATLNSLENOIKN 20

RESULT 10
ID ABO69131 standard; protein; 571 AA.
XX ABO69131;
XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #1306.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
XX US6551795-B1.
XX 22-APR-2003.
XX 18-FEB-1999; 99US-0025291.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX N-PSDB; ABD02702.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 17877; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polymucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant

CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84336 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX
SQ Sequence 571 AA;

Query Match 9.3%; Score 92; DB 7; Length 571;
Best Local Similarity 21.6%; Pred. No. 2.7;
Matches 42; Conservative 24; Mismatches 54; Indels 74; Gaps 9;

QY 15 ENOINSGSGSTA--SPGIFRYVDCG---YARAHK-----MROILMNGYDC 60
Db 294 FTQVR--SCGISTAVSPGMSQYFREDYSDKAKTHGGLDILORAGVQVIMLENSDC 351
QY 61 EKQFYGNLKAATGTC-----CVMSYHVALIVSYRNAGSVTER 99
Db 352 K-----GTCRLVRPNRDIPTQSPFCGDKNCLDESLVGL--GEYIDGLQDD 396
QY 100 RII-----DPSLFSSGPEVTDTAMRNACVNTSCGSASVSSYANTAGVYTRFSNSTL 151
Db 397 AIIVLHSDGSHGPEYERYPKEMERFQPVCRITNOLGSCSKELVN-----V 442
QY 152 YDNNLINTNCVLT/K 165
Db 443 YDNTILYTDHFLTK 456

RESULT 11
ID AAG52479
XX AAG52479 standard; protein; 594 AA.
XX AAG52479;
XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 66714.
XX Arabidopsis thaliana.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-014053P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141642P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-015138P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153707P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158025P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 03-AUG-1999; 99US-0146389P.
 PR 04-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 05-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 06-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 09-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 10-AUG-1999; 99US-0147935P.
 PR 11-AUG-1999; 99US-0148171P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149802P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-015138P.
 PR 01-SEP-1999; 99US-015130P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157855P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158392P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 18-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.

PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161923P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 9.2%; Score 91; DB 3; Length 594;
 Best Local Similarity 24.6%; Pred. No. 3.7;
 Matches 31; Conservative 14; Mismatches 45; Indels 36; Gaps 5;

QY 25 TSTASPCITFRY---PVD--GCYARAHKMQILMNGY---DCKQFYGNLAKST-- 73
 Db 442 TIVDDNPSTVFYFPHVPLQRCVEAIRLVSKVTSNRPFLNTQCCKQVHQLSLSVKA 501

QY 74 -----GTCVWASYHVALVS-----YKNASGVTEKRIIDPSLFS 108
 Db 502 NINLRPKQNDTKSMAQFCQDVIWVITWYHGGCLVGVKVPARKVLGVDRRLAVIDGSTD 561

QY 109 GGPVTD 114
 Db 562 EBPQTN 567

RESULT 13
 ABB91568
 ID ABB91568 strand; protein; 594 AA.
 XX
 AC ABB91568;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE Herbicidally active polypeptide SEQ ID NO 779.
 XX
 KW Herbicidal; plant; agriculture; herbicide.
 OS Arabidopsis thaliana.
 PN WO200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 2001WO-EP009892.
 XX
 PR 28-AUG-2001; 2001WO-EP009892.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Tietjen K, Weidier M;
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.
 XX
 PS Claim 5; SEQ ID NO 779; 261bp + Sequence Listing; English.
 XX
 CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search

XXV

55

PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI, 2004-061375/06.

XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

PS Claim 1, SEQ ID NO 6569; 122bp; English.

XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned
CC to provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 927 AA;

Query Match 9.0%; Score 89.5; DB 8; Length 927;

Best Local Similarity 20.1%; Pred. No. 9.6; Matches 61; Conservative 25; Mismatches 68; Indels 149; Gaps 16;

QY 4 VI PDVATLNSLRFNQI--KNOSCGSTASSPCITFRYPVDGCVARAHKMQI-----LMN 55
DB 130 VCPDYASGRFCQNEIKKDNKSCGKNA-----DCYVANHQLNCKCKPGYTARR 176
QY 56 NGYDEKQ-----FVGNLKAS-TGTCVANS----- 81
DB 177 NGRDDCMKVKQACMSGDPRHYVTYDGLRFPYQGTCPYVFSQPCITLPAFLWMSVRAKNEL 236
QY 82 ---YHVA---ILVSYKN-----ASGVTEKRIIDPSIFS-----S 109
DB 237 PKGYHISQVSEVEVDLHNLITHVDRSKTALVNGV---QVLTPMWFPMKNTWTVVRFS 293
QY 110 GP-----YDITARNACV-----NTSGSA-----SVSSYANTAGNV 141
DB 294 GSTFTIENDQGVVYFTTNSLCCVOVPDIPERNGATTLGCLAGNIDGKLDVNVNNGSV 353
QY 142 Y-----YRSPSN-----SYLDNNLI-----NTNCVLTKEFSLLS 170
DB 354 LAIKSRQPENNHNDPMFKETDPTWITDKFLILRPGQENCINGOTLDNNTNCVSTISILAQ 413
QY 171 GCS 173
DB 414 SCA 416

Search completed: September 16, 2006, 06:18:16
Job time : 201 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 06:23:35; Search time 50 Seconds
(without alignments)
323.865 Million cell updates/sec

Title: US-10-815-774-6
Perfect score: 991
Sequence: 1 LASV1PDVATLNSLFNOJKN.....FSLSGCSPAPDVSSCGF 185

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /EMC Celerra_SIDS3/ptodata/2/1aa/5_COMB.pep:*
2: /EMC Celerra_SIDS3/ptodata/2/1aa/6_COMB.pep:*
3: /EMC Celerra_SIDS3/ptodata/2/1aa/7_COMB.pep:*
4: /EMC Celerra_SIDS3/ptodata/2/1aa/H_COMB.pep:*
5: /EMC Celerra_SIDS3/ptodata/2/1aa/PTOS_COMB.pep:*
6: /EMC Celerra_SIDS3/ptodata/2/1aa/RE_COMB.pep:*
7: /EMC Celerra_SIDS3/ptodata/2/1aa/backflist.pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	991	100.0	185	2 US-09-727-769A-6	Sequence 6, Appl
2	991	100.0	320	2 US-09-727-769A-8	Sequence 8, Appl
3	874	88.2	185	2 US-09-324-910-6	Sequence 6, Appl
4	874	88.2	185	2 US-09-793-495-6	Sequence 6, Appl
5	874	88.2	319	2 US-09-324-910-12	Sequence 12, Appl
6	874	88.2	319	2 US-09-793-495-12	Sequence 12, Appl
7	108	10.9	20	2 US-09-727-769A-2	Sequence 1, Appl
8	96	9.7	20	2 US-09-727-769A-1	Sequence 1, Appl
9	92	9.3	571	2 US-09-252-540-12607	Sequence 12607, A
10	83.5	8.4	459	2 US-09-324-910-2	Sequence 2, Appl
11	83	8.4	20	2 US-09-793-495-2	Sequence 2, Appl
12	83	8.4	20	2 US-09-793-495-2	Sequence 2, Appl
13	82	8.3	927	2 US-09-198-452A-472	Sequence 472, App
14	82	8.3	937	2 US-09-438-185A-449	Sequence 449, App
15	80.5	8.1	710	2 US-09-248-796A-14136	Sequence 14136, A
16	80	8.1	406	2 US-09-248-796A-14345	Sequence 14345, A
17	79	8.0	20	2 US-09-324-910-1	Sequence 1, Appl
18	79	8.0	20	2 US-09-793-495-1	Sequence 1, Appl
19	78	7.9	234	2 US-09-252-991A-29027	Sequence 29027, A
20	78	7.9	357	2 US-09-270-767-45129	Sequence 45129, A
21	77.5	7.8	666	2 US-08-982-785A-11	Sequence 11, Appl
22	77.5	7.8	666	2 US-09-629-498-11	Sequence 11, Appl
23	76.5	7.7	433	2 US-09-104-623A-4	Sequence 4, Appl
24	76.5	7.7	433	2 US-09-019-532-4	Sequence 4, Appl
25	76.5	7.7	433	2 US-09-338-746-4	Sequence 4, Appl
26	76.5	7.7	433	2 US-09-985-689A-3	Sequence 3, Appl

27	76.5	7.7	433	2 US-09-985-689A-4	Sequence 4, Appl
28	76.5	7.7	433	2 US-09-985-689A-5	Sequence 5, Appl
29	76.5	7.7	499	1 US-07-952-853-6	Sequence 6, Appl
30	76.5	7.7	499	1 US-08-914-848-6	Sequence 6, Appl
31	76.5	7.7	635	1 US-08-873-479-43	Sequence 43, Appl
32	76.5	7.7	726	2 US-09-949-016-5916	Sequence 5916, Ap
33	76.5	7.7	737	2 US-09-949-016-11607	Sequence 11607, A
34	76	7.7	1133	2 US-09-902-540-12243	Sequence 12243, A
35	75.5	7.6	149	2 US-09-489-039A-8252	Sequence 8252, Ap
36	75.5	7.6	468	3 US-09-497-967-54	Sequence 54, Appl
37	75.5	7.6	539	2 US-09-902-540-11983	Sequence 11983, A
38	75.5	7.6	1190	2 US-09-949-016-7676	Sequence 7676, Ap
39	75.5	7.6	1190	2 US-09-949-016-7677	Sequence 7677, Ap
40	75.5	7.6	1219	2 US-10-148-806-2	Sequence 2, Appl
41	75	7.6	271	1 US-08-467-155A-10	Sequence 10, Appl
42	75	7.6	271	1 US-08-628-198-10	Sequence 10, Appl
43	75	7.6	271	2 US-09-201-038-10	Sequence 10, Appl
44	75	7.6	271	3 US-10-021-368-10	Sequence 10, Appl
45	75	7.6	271	5 PCT-US96-07343-10	Sequence 10, Appl

ALIGNMENTS

```
RESULT 1
US-09-727-769A-6
Sequence 6, Application US/09727769A
Patent No. 6756221
GENERAL INFORMATION:
APPLICANT: Amaro Enzyme, Inc.
TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SA
FILE REFERENCE: 062106
CURRENT FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: JP 11-345044
PRIOR FILING DATE: 1999-12-03
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
TYPE: PRT
LENGTH: 185
ORGANISM: Cryobacterium sp. No. 6756221 9670
US-09-727-769A-6

Query Match      100.0%; Score 991; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.6e-101;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASV1PDVATLNSLFNOJKNKOSCGTSTASPCITFRYPVDCGYARAHKROILMNGYDC 60
DB 1 LASV1PDVATLNSLFNOJKNKOSCGTSTASPCITFRYPVDCGYARAHKROILMNGYDC 60
QY 61 EKQFYGNLKAISTGTCCVAMSYHVALIVSYKASGVTEKRIIDPSLFSGGPYDTAMRNA 120
DB 61 EKQFYGNLKAISTGTCCVAMSYHVALIVSYKASGVTEKRIIDPSLFSGGPYDTAMRNA 120
QY 121 CNTVSCGASVSANTAGNYRRSPNSYILNDNLINTNCTLTFTSLSGCSPAPDV 180
DB 121 CNTVSCGASVSANTAGNYRRSPNSYILNDNLINTNCTLTFTSLSGCSPAPDV 180
QY 181 SSCGF 185
DB 181 SSCGF 185

RESULT 2
US-09-727-769A-8
Sequence 8, Application US/09727769A
Patent No. 6756221
GENERAL INFORMATION:
APPLICANT: Amaro Enzyme, Inc.
TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SA
```

```
/ TITLE OF INVENTION: ENCODING THE SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
/ FILE REFERENCE: 062106
/ CURRENT APPLICATION NUMBER: US/09/727,769A
/ CURRENT FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: JP 11-345044
/ PRIOR FILING DATE: 1999-12-03
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 8
/ LENGTH: 320
/ TYPE: PRF
/ ORGANISM: Chryseobacterium sp. No. 6756221 9670
US-09-727-769A-8
```

```
Query Match          100.0%; Score 991; DB 2; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.2e-100;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LASVDPVATLNSLFNQIKNOSCGTSTASPCITFRYPVDCGYARAHKROILMNNGYDC 60
DB 136 LASVDPVATLNSLFNQIKNOSCGTSTASPCITFRYPVDCGYARAHKROILMNNGYDC 195
QY 61 EKQFYGNLKASTGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSGPGVTDAMRNA 120
DB 196 EKQFYGNLKASTGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSGPGVTDAMRNA 255
QY 121 CVNTSCGSASVSYANTAGNYYRSPNSLYLDNNLINTNCVLTKEFSLSGCSPSPADPV 180
DB 256 CVNTSCGSASVSYANTAGNYYRSPNSLYLDNNLINTNCVLTKEFSLSGCSPSPADPV 315
QY 181 SSCGF 185
DB 316 SSCGF 320
```

```
RESULT 3
US-09-324-910-6
/ Sequence 6, Application US/09324910
/ Patent No. 6251651
/ GENERAL INFORMATION:
/ APPLICANT: Yamaguchi, Shotaro
/ APPLICANT: Matsura, Akira
/ TITLE OF INVENTION: NOVEL, PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE
/ TITLE OF INVENTION: SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
/ FILE REFERENCE: 0-54362
/ CURRENT APPLICATION NUMBER: US/09/324,910
/ CURRENT FILING DATE: 1999-06-03
/ EARLIER APPLICATION NUMBER: HEI-10-173940
/ EARLIER FILING DATE: 1998-06-04
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 185
/ TYPE: PRF
/ ORGANISM: Chryseobacterium gleum
US-09-324-910-6
```

```
Query Match          88.2%; Score 874; DB 2; Length 185;
Best Local Similarity 87.4%; Pred. No. 4.3e-88;
Matches 160; Conservative 10; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 3 SVIPDVALTNSLFNQIKNOSCGTSTASPCITFRYPVDCGYARAHKROILMNNGYDC 62
DB 3 SVIPDVALTNSLFNQIKNOSCGTSTASPCITFRYPVDCGYARAHKROILMNNGYDC 62
QY 63 QFYGNLKASTGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSGPGVTDAMRNA 122
DB 63 QFYGNLKASTGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSGPGVTDAMRNA 122
QY 123 NTS CGSASVSYANTAGNYYRSPNSLYLDNNLINTNCVLTKEFSLSGCSPSPADPV 182
DB 123 NTS CGSASVSYANTAGNYYRSPNSLYLDNNLINTNCVLTKEFSLSGCSPSPADPV 182
```

```
QY 183 CGF 185
DB 183 CGF 185
```

```
RESULT 4
US-09-793-495-6
/ Sequence 6, Application US/09793495
/ Patent No. 6770469
/ GENERAL INFORMATION:
/ APPLICANT: Yamaguchi, Shotaro
/ APPLICANT: Matsura, Akira
/ TITLE OF INVENTION: NOVEL, PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE
/ TITLE OF INVENTION: SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
/ FILE REFERENCE: 0-54362
/ CURRENT APPLICATION NUMBER: US/09/793,495
/ CURRENT FILING DATE: 2001-02-27
/ PRIOR APPLICATION NUMBER: 09/324,910
/ PRIOR FILING DATE: 1999-06-03
/ PRIOR APPLICATION NUMBER: HEI-10-173940
/ PRIOR FILING DATE: 1998-06-04
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 185
/ TYPE: PRF
/ ORGANISM: Chryseobacterium gleum
US-09-793-495-6
```

```
Query Match          88.2%; Score 874; DB 2; Length 185;
Best Local Similarity 87.4%; Pred. No. 4.3e-88;
Matches 160; Conservative 10; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 3 SVIPDVALTNSLFNQIKNOSCGTSTASPCITFRYPVDCGYARAHKROILMNNGYDC 62
DB 3 SVIPDVALTNSLFNQIKNOSCGTSTASPCITFRYPVDCGYARAHKROILMNNGYDC 62
QY 63 QFYGNLKASTGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSGPGVTDAMRNA 122
DB 63 QFYGNLKASTGTCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSGPGVTDAMRNA 122
QY 123 NTS CGSASVSYANTAGNYYRSPNSLYLDNNLINTNCVLTKEFSLSGCSPSPADPV 182
DB 123 NTS CGSASVSYANTAGNYYRSPNSLYLDNNLINTNCVLTKEFSLSGCSPSPADPV 182
QY 183 CGF 185
DB 183 CGF 185
```

```
RESULT 5
US-09-324-910-12
/ Sequence 12, Application US/09324910
/ Patent No. 6251651
/ GENERAL INFORMATION:
/ APPLICANT: Yamaguchi, Shotaro
/ APPLICANT: Matsura, Akira
/ TITLE OF INVENTION: NOVEL, PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE
/ TITLE OF INVENTION: SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
/ FILE REFERENCE: 0-54362
/ CURRENT APPLICATION NUMBER: US/09/324,910
/ CURRENT FILING DATE: 1999-06-03
/ EARLIER APPLICATION NUMBER: HEI-10-173940
/ EARLIER FILING DATE: 1998-06-04
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 12
/ LENGTH: 319
/ TYPE: PRF
/ ORGANISM: Chryseobacterium gleum
US-09-324-910-12
```

```
Query Match          88.2%; Score 874; DB 2; Length 319;
```



```

Query Match          9.3%; Score 92; DB 2; Length 571;
Best Local Similarity 21.6%; Pred. No. 0.29;
Matches 42; Conservative 24; Mismatches 54; Indels 74; Gaps 9;

QY 15 ENQIKNSGCTGA--SSPITRRYPVDC---YARAHK-----MROILMNGVYC 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 294 FTQVR--SCGTSTAVSVPCMFQSPREDYSDKAKTHBLLDLORAGVQVIMLENSDC 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 EKQFVYGNLKAATGTC-----CVAMSYHVALIVSYRNAGSVTEK 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 352 K-----GTLARVPNDIPKTOPSPFCGKNCCLDESLVGL---QEVIGLQGD 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 RII-----DPSLFSGGPVTDITAMRNACVNTSCGSASVSYANTAGNVYRSPNSYL 151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 397 AIIVLHSDSGHGEYERYPKEMERFQPVCRITQLGSCSKELIVN-----V 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 152 YDNNLINTNCVLT 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 443 YDNTILYTDHFLTK 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-09-902-540-12607
; Sequence 12607, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12607
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12607

Query Match          8.4%; Score 83.5; DB 2; Length 459;
Best Local Similarity 23.4%; Pred. No. 1.8;
Matches 47; Conservative 21; Mismatches 60; Indels 73; Gaps 12;

QY 2 ASYTPDVATINSLFN-----QIKNSGCTSTASSP-CIT---FRYPVDG 41
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 226 ALVVEDVINANALPNTGSGMPEBSGHTFLYQVR---YTIASVPLCPGANWAVPVKG 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 42 CYAAHMRQIIMNNGVCEKQFYGNLKAATGTCVAMSYHVALIVSYRNAGSVTEKRI 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 281 LFD-----LOGDRYESPRAFTLG--CDAGVIAKCYRWGHPWL-----DOAT----- 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 102 IDBSLFSSGPTVDTAMRNACVNTSCGSASVSYANTAGNVYRSPNSYLDNNLINTNC 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 321 -----SGPITTEAHM-----SCTMRADYC-----GQGTSTFTLDGTRLRPMD 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 162 VITKFSLLSGSPAPDVSS 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 358 ALTP-AIIS---APSPDSSS 373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-09-324-910-2
; Sequence 2, Application US/09324910
; Patent No. 6251651
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; APPLICANT: Matsuura, Akira
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE

```

```

; TITLE OF INVENTION: SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
; FILE REFERENCE: Q-54362
; CURRENT APPLICATION NUMBER: US/09/324,910
; CURRENT FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: HEI-10-173940
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Chryseobacterium gleum
US-09-324-910-2

Query Match          8.4%; Score 83; DB 2; Length 20;
Best Local Similarity 78.9%; Pred. No. 0.024;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 145 SPSNSYLDNNLINTNCVL 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 SPSSGLYDNNYVNTNCVL 19
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-09-793-495-2
; Sequence 2, Application US/09793495
; Patent No. 6770469
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; APPLICANT: Matsuura, Akira
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE
; TITLE OF INVENTION: SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
; FILE REFERENCE: Q-54362
; CURRENT APPLICATION NUMBER: US/09/793,495
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/324,910
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: HEI-10-173940
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Chryseobacterium gleum
US-09-793-495-2

Query Match          8.4%; Score 83; DB 2; Length 20;
Best Local Similarity 78.9%; Pred. No. 0.024;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 145 SPSNSYLDNNLINTNCVL 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 SPSSGLYDNNYVNTNCVL 19
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-198-452A-472
; Sequence 472, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 472
; LENGTH: 927
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae

```


This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 06:35:10 ; Search time 173 Seconds

(without alignments)
495.345 Million cell updates/sec

Title: US-10-815-774-6

Perfect score: 991

Sequence: 1 LASVIPDVATINSLEFNOIKN.....FSLISGCSPPAPDVSSCGF 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues 2097797

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	185	4	US-10-815-774-6
2	991	100.0	320	4	US-10-815-774-8
3	874	88.2	185	3	US-09-793-495-6
4	874	88.2	185	4	US-10-815-751-6
5	874	88.2	319	3	US-09-793-495-12
6	874	88.2	319	4	US-10-815-751-12
7	108	10.9	20	4	US-10-815-774-2
8	96	9.7	20	4	US-10-815-774-1
9	95	9.6	1351	6	US-11-129-741-2937
10	95	9.6	1385	6	US-11-129-741-3655
11	93	9.4	1351	6	US-11-129-741-2947
12	90.5	9.1	276	4	US-10-425-115-197136
13	90.5	9.1	292	4	US-10-425-114-64881
14	89.5	9.0	927	4	US-10-369-493-6569
15	89.5	9.0	927	4	US-10-369-493-6571
16	89	9.0	1895	4	US-10-399-136-3
17	86	8.7	1219	4	US-10-341-200-23
18	85.5	8.6	442	4	US-10-425-114-41827
19	85.5	8.6	477	6	US-11-004-418A-34
20	85.5	8.6	477	6	US-11-089-551A-34
21	85.5	8.6	961	4	US-10-425-114-71401
22	85.5	8.6	1142	4	US-10-425-114-62440
23	85.5	8.6	1171	4	US-10-425-115-360704
24	84	8.5	881	4	US-10-437-963-111609
25	84	8.5	3726	6	US-11-097-143-18633
26	83	8.4	20	3	US-09-793-495-2
27	83	8.4	20	4	US-10-815-751-2

28	83	8.4	1493	4	US-10-437-963-115870	Sequence 115870,
29	82.5	8.3	5374	4	US-10-028-248A-75	Sequence 75, Appl
30	82.5	8.3	5374	4	US-10-107-782-75	Sequence 75, Appl
31	82.5	8.3	5376	4	US-10-028-248A-74	Sequence 74, Appl
32	82.5	8.3	5376	4	US-10-107-782-74	Sequence 74, Appl
33	82	8.3	703	4	US-10-369-493-22435	Sequence 22435, A
34	82	8.3	927	4	US-10-289-762-472	Sequence 472, App
35	82	8.3	928	4	US-10-312-273-33	Sequence 33, Appl
36	82	8.3	928	5	US-10-503-135-99	Sequence 99, Appl
37	81	8.2	346	6	US-11-097-143-41370	Sequence 41370, A
38	81	8.2	839	6	US-11-188-298-21025	Sequence 21025, A
39	79.5	8.0	1894	4	US-10-087-684-81	Sequence 81, Appl
40	79.5	8.0	1894	4	US-10-218-779-81	Sequence 81, Appl
41	79	8.0	20	3	US-09-793-495-1	Sequence 1, Appl
42	79	8.0	20	4	US-10-815-751-1	Sequence 1, Appl
43	79	8.0	221	6	US-11-087-099-2126	Sequence 2126, Ap
44	79	8.0	791	5	US-10-469-204-146	Sequence 146, App
45	79	8.0	1290	6	US-11-097-143-15246	Sequence 15246, A

ALIGNMENTS

```

RESULT 1
US-10-815-774-6
; Sequence 6, Application US/10815774
; Publication No. US20040175799A1
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
; FILE REFERENCE: 080844
; CURRENT APPLICATION NUMBER: US/10/815,774
; PRIOR FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US/09/727,769
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US/09/324,910
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: JP Hei. 11-345044
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP Hei. 10-173940
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Cryseobacterium sp. No. 9670
US-10-815-774-6
Query Match
Best Local Similarity 100.0%; Pred. No. 2.3e-92;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 LASVIPDVATINSLEFNOIKNSCGTSTASPCITFRYPVDCGYAHAKKROJIMNNGYDC 60
1 LASVIPDVATINSLEFNOIKNSCGTSTASPCITFRYPVDCGYAHAKKROJIMNNGYDC 60
61 EKQFYGNLKAISTGTCCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSSGPTDTAMRNA 120
61 EKQFYGNLKAISTGTCCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSSGPTDTAMRNA 120
61 EKQFYGNLKAISTGTCCVAMSYHVALIVSYKNASGVTEKRIIDPSLFSSGPTDTAMRNA 120
121 CMTSCGASAVSYANTGNYRSPNSYLYDNLIMTNCVLTPEFSLISGSPAPDV 180
121 CMTSCGASAVSYANTGNYRSPNSYLYDNLIMTNCVLTPEFSLISGSPAPDV 180
121 CMTSCGASAVSYANTGNYRSPNSYLYDNLIMTNCVLTPEFSLISGSPAPDV 180
QY 181 SSCGF 185
QY 181 SSCGF 185
Db 181 SSCGF 185
Db 181 SSCGF 185
RESULT 2
US-10-815-774-8

```

```
; Sequence 8, Application US/10815774
; Publication No. US20040175799A1
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
; FILE REFERENCE: 080844
; CURRENT APPLICATION NUMBER: US/10/815,774
; PRIORITY FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US/09/727,769
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US/09/324,910
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: JP Hei. 11-345044
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP Hei. 10-173940
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Cryseobacterium sp. No. 9670
US-10-815-774-6

Query Match          100.0%; Score 991; DB 4; Length 320;
Best Local Similarity 100.0%; Pred. No. 4,6e-92;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASVDPVATLNSLFTQIKNOSCGTSTASSPCITFRYPVDCYARAHKROILLMNGYDC 60
    |||||
DB 136 LASVDPVATLNSLFTQIKNOSCGTSTASSPCITFRYPVDCYARAHKROILLMNGYDC 195
    |||||

QY 61 EKQFVGNLKASTGTCCVAMSYHVALIVSYKNSGVTGKRIIDPSLFSGSPVDTAMRNA 120
    |||||
DB 196 EKQFVGNLKASTGTCCVAMSYHVALIVSYKNSGVTGKRIIDPSLFSGSPVDTAMRNA 255
    |||||

QY 121 CVNTSCGSASVSYANTAGNVYRSPNSYLYDNNLINTNCVLTGKFSLSGCSPPAPDV 180
    |||||
DB 256 CVNTSCGSASVSYANTAGNVYRSPNSYLYDNNLINTNCVLTGKFSLSGCSPPAPDV 315
    |||||

QY 181 SSCGF 185
    |||||
DB 316 SSCGF 320

RESULT 3
US-09-793-495-6
; Sequence 6, Application US/09793495
; Publication No. US20040072318A1
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
; FILE REFERENCE: 0-54362
; CURRENT APPLICATION NUMBER: US/09/793,495
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/324,910
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: HEI-10-173940
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Chryseobacterium gleum
US-09-793-495-6

Query Match          88.2%; Score 874; DB 3; Length 185;
Best Local Similarity 87.4%; Pred. No. 1.8e-80;
Matches 160; Conservative 10; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 3 SVIPDVATLNSLFTQIKNOSCGTSTASSPCITFRYPVDCYARAHKROILLMNGYDC 62
    |||||
DB 3 SVIPDVATLNSLFTQIKNOSCGTSTASSPCITFRYPVDCYARAHKROILLMNGYDC 62
    |||||

QY 63 QFVGNLKASTGTCCVAMSYHVALIVSYKNSGVTGKRIIDPSLFSGSPVDTAMRNA 122
    |||||
DB 63 QFVGNLKASTGTCCVAMSYHVALIVSYKNSGVTGKRIIDPSLFSGSPVDTAMRNA 122
    |||||

QY 123 NTS CGSASVSYANTAGNVYRSPNSYLYDNNLINTNCVLTGKFSLSGCSPPAPDV 182
    |||||
DB 123 NTS CGSASVSYANTAGNVYRSPNSYLYDNNLINTNCVLTGKFSLSGCSPPAPDV 182
    |||||

QY 183 CGF 185
    |||||
DB 183 CGF 185

RESULT 4
US-10-815-751-6
; Sequence 6, Application US/10815751
; Publication No. US2004016558A1
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
; FILE REFERENCE: 0-54362
; CURRENT APPLICATION NUMBER: US/10/815,751
; PRIOR FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US/09/793,495
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/324,910
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: HEI-10-173940
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Chryseobacterium gleum
US-10-815-751-6

Query Match          88.2%; Score 874; DB 4; Length 185;
Best Local Similarity 87.4%; Pred. No. 1.8e-80;
Matches 160; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 3 SVIPDVATLNSLFTQIKNOSCGTSTASSPCITFRYPVDCYARAHKROILLMNGYDC 62
    |||||
DB 3 SVIPDVATLNSLFTQIKNOSCGTSTASSPCITFRYPVDCYARAHKROILLMNGYDC 62
    |||||

QY 63 QFVGNLKASTGTCCVAMSYHVALIVSYKNSGVTGKRIIDPSLFSGSPVDTAMRNA 122
    |||||
DB 63 QFVGNLKASTGTCCVAMSYHVALIVSYKNSGVTGKRIIDPSLFSGSPVDTAMRNA 122
    |||||

QY 123 NTS CGSASVSYANTAGNVYRSPNSYLYDNNLINTNCVLTGKFSLSGCSPPAPDV 182
    |||||
DB 123 NTS CGSASVSYANTAGNVYRSPNSYLYDNNLINTNCVLTGKFSLSGCSPPAPDV 182
    |||||

QY 183 CGF 185
    |||||
DB 183 CGF 185

RESULT 5
US-09-793-495-12
; Sequence 12, Application US/09793495
; Publication No. US20040072318A1
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, Shotaro
; TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE
```

1 TITLE OF INVENTION: SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
2 FILE REFERENCE: Q-54362
3 CURRENT APPLICATION NUMBER: US/09/793,495
4 CURRENT FILING DATE: 2001-02-27
5 PRIOR APPLICATION NUMBER: 09/324,910
6 PRIOR FILING DATE: 1999-06-03
7 PRIOR APPLICATION NUMBER: HEI-10-173940
8 PRIOR FILING DATE: 1998-06-04
9 NUMBER OF SEQ ID NOS: 12
10 SOFTWARE: PatentIn Ver. 2.0
11 SEQ ID NO 12
12 LENGTH: 319
13 TYPE: PRT
14 ORGANISM: *Chryseobacterium gleum*
15 US-09-793-495-12

Query Match 88.2%; Score 874; DB 3; Length 319;
Best Local Similarity 87.4%; Pred. No. 3, 6e-80;
Matches 160; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 3 SVTPVATLNSL.FNOIKNOSCGSTSPCITFRYPVDCYARAHKROILLNNGYDCEK 62
DB 13 SVTPDATTNLSLFTQIKNOACGTSTASPCITFRYPVDCYARAHKROILLNAGYDCEK 196
QY 63 QFVYGNLKASTGTCCVAMSVHYAIIIVSYKNASGVTEKRIIDPSLFSSGPFVTDANRNACV 122
DB 197 QFVYGNLRASTGTCCVAMSVHYAIIIVSYKNASGIVEKRIIDPSLFSSGPFVTDANRNAACT 256
QY 123 NTSCGSASVSYANTAGNVYRSPNSNYLYDNNLINTNCVLTFRSLSGCSPSPADVSS 182
DB 257 NTSCGSASVSYANTAGNVYRSPSGSLLYDNNYNTNCVLTNIFSLSGCSPSPAPSVAS 316
QY 183 CGF 185
DB 317 CGF 319

RESULT 6
US-10-815-751-12
1 Sequence 12, Application US/10815751
2 Publication No. US2004016658A1
3 GENERAL INFORMATION:
4 APPLICANT: Yamaguchi, Shotaro
5 TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, GENE ENCODING THE
6 TITLE OF INVENTION: SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
7 FILE REFERENCE: Q-54362
8 CURRENT APPLICATION NUMBER: US/10/815,751
9 CURRENT FILING DATE: 2004-04-02
10 PRIOR APPLICATION NUMBER: US/09/793,495
11 PRIOR FILING DATE: 2001-02-27
12 PRIOR APPLICATION NUMBER: 09/324,910
13 PRIOR FILING DATE: 1999-06-03
14 PRIOR APPLICATION NUMBER: HEI-10-173940
15 PRIOR FILING DATE: 1998-06-04
16 NUMBER OF SEQ ID NOS: 12
17 SOFTWARE: PatentIn Ver. 2.0
18 SEQ ID NO 12
19 LENGTH: 319
20 TYPE: PRT
21 ORGANISM: *Chryseobacterium gleum*
22 US-10-815-751-12

Query Match 88.2%; Score 874; DB 4; Length 319;
Best Local Similarity 87.4%; Pred. No. 3, 6e-80;
Matches 160; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 3 SVTPVATLNSL.FNOIKNOSCGSTSPCITFRYPVDCYARAHKROILLNNGYDCEK 62
DB 13 SVTPDATTNLSLFTQIKNOACGTSTASPCITFRYPVDCYARAHKROILLNAGYDCEK 196
QY 63 QFVYGNLKASTGTCCVAMSVHYAIIIVSYKNASGVTEKRIIDPSLFSSGPFVTDANRNACV 122

DB 197 QFVYGNLRASTGTCCVAMSVHYAIIIVSYKNASGIVEKRIIDPSLFSSGPFVTDANRNAACT 256
QY 123 NTSCGSASVSYANTAGNVYRSPNSNYLYDNNLINTNCVLTFRSLSGCSPSPADVSS 182
DB 257 NTSCGSASVSYANTAGNVYRSPSGSLLYDNNYNTNCVLTNIFSLSGCSPSPAPSVAS 316
QY 183 CGF 185
DB 317 CGF 319

RESULT 7
US-10-815-774-2
1 Sequence 2, Application US/10815774
2 Publication No. US2004015799A1
3 GENERAL INFORMATION:
4 APPLICANT: Yamaguchi, Shotaro
5 TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SAM
6 TITLE OF INVENTION: ENCODING THE SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
7 FILE REFERENCE: Q80844
8 CURRENT APPLICATION NUMBER: US/10/815,774
9 CURRENT FILING DATE: 2004-04-02
10 PRIOR APPLICATION NUMBER: US/09/727,769
11 PRIOR FILING DATE: 2000-12-04
12 PRIOR APPLICATION NUMBER: US/09/324,910
13 PRIOR FILING DATE: 1999-06-03
14 PRIOR APPLICATION NUMBER: JP Hei. 11-345044
15 PRIOR FILING DATE: 1999-12-03
16 PRIOR APPLICATION NUMBER: JP Hei. 10-173940
17 PRIOR FILING DATE: 1998-06-04
18 NUMBER OF SEQ ID NOS: 11
19 SOFTWARE: PatentIn version 3.1
20 SEQ ID NO 2
21 LENGTH: 20
22 TYPE: PRT
23 ORGANISM: *Chryseobacterium sp.* No. 9670
24 US-10-815-774-2

Query Match 10.9%; Score 108; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 SPSNSLYDNNLINTNCVLT 164
DB 1 SPSNSLYDNNLINTNCVLT 20

RESULT 8
US-10-815-774-1
1 Sequence 1, Application US/10815774
2 Publication No. US20040175799A1
3 GENERAL INFORMATION:
4 APPLICANT: Yamaguchi, Shotaro
5 TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SAM
6 TITLE OF INVENTION: ENCODING THE SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF
7 FILE REFERENCE: Q80844
8 CURRENT APPLICATION NUMBER: US/10/815,774
9 CURRENT FILING DATE: 2004-04-02
10 PRIOR APPLICATION NUMBER: US/09/727,769
11 PRIOR FILING DATE: 2000-12-04
12 PRIOR APPLICATION NUMBER: US/09/324,910
13 PRIOR FILING DATE: 1999-06-03
14 PRIOR APPLICATION NUMBER: JP Hei. 11-345044
15 PRIOR FILING DATE: 1999-12-03
16 PRIOR APPLICATION NUMBER: JP Hei. 10-173940
17 PRIOR FILING DATE: 1998-06-04
18 NUMBER OF SEQ ID NOS: 11
19 SOFTWARE: PatentIn version 3.1
20 SEQ ID NO 1
21 LENGTH: 20
22 TYPE: PRT
23 ORGANISM: *Chryseobacterium sp.* No. 9670
24 US-10-815-774-1


```
APPLICANT: Alexandrov, Nickolai
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE OF INVENTION: Polypeptides Encoded Thereby
FILE REFERENCE: 2750-1590PUS2
CURRENT APPLICATION NUMBER: US/11/056,355B
PRIOR FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: 60/544,190
NUMBER OF SEQ ID NOS: 119966
LENGTH: 594
TYPE: prt
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(594)
OTHER INFORMATION: Ceres Seq. ID no. 12668218
US-11-056-355B-83015
```

Query Match
Best Local Similarity 24.6%; Pred. No. 0.87;
Matches 31; Conservative 14; Mismatches 45; Indels 36; Gaps 5;

```
QY 25 TSTASPCITFRY---PVD--GCYARAHKROILMNGY---DCEKOFYGNLKAST-- 73
DB 442 TNVDNPSVTFFNYFKHPVDLQRCVEAIRLVSKVTSNRPLNTQCCKQNVHKLSTSVKA 501
QY 74 -----GTCCVMSYHVALIVS-----YKNASGVTEKRIIDPSLFS 108
DB 502 NINLRPKQNDTKSMAQFCSDTVTTIMHYHGCLVGVKVSPPNRKVLGVDRLRVIDGSTFD 561
QY 109 SGPVTD 114
DB 562 ESRGTN 567
```

RESULT 3

```
US-11-056-355B-91487
; Sequence 91487, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 91487
; LENGTH: 594
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(594)
; OTHER INFORMATION: Ceres Seq. ID no. 12727965
US-11-056-355B-91487
```

Query Match
Best Local Similarity 24.6%; Pred. No. 0.87;
Matches 31; Conservative 14; Mismatches 45; Indels 36; Gaps 5;

```
QY 25 TSTASPCITFRY---PVD--GCYARAHKROILMNGY---DCEKOFYGNLKAST-- 73
DB 442 TNVDNPSVTFFNYFKHPVDLQRCVEAIRLVSKVTSNRPLNTQCCKQNVHKLSTSVKA 501
QY 74 -----GTCCVMSYHVALIVS-----YKNASGVTEKRIIDPSLFS 108
DB 502 NINLRPKQNDTKSMAQFCSDTVTTIMHYHGCLVGVKVSPPNRKVLGVDRLRVIDGSTFD 561
```

```
QY 109 SGPVTD 114
DB 562 ESRGTN 567
```

RESULT 4

```
US-11-056-355B-95243
; Sequence 95243, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 95243
; LENGTH: 594
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(594)
; OTHER INFORMATION: Ceres Seq. ID no. 12727965
US-11-056-355B-95243
```

Query Match
Best Local Similarity 24.6%; Pred. No. 0.87;
Matches 31; Conservative 14; Mismatches 45; Indels 36; Gaps 5;

```
QY 25 TSTASPCITFRY---PVD--GCYARAHKROILMNGY---DCEKOFYGNLKAST-- 73
DB 442 TNVDNPSVTFFNYFKHPVDLQRCVEAIRLVSKVTSNRPLNTQCCKQNVHKLSTSVKA 501
QY 74 -----GTCCVMSYHVALIVS-----YKNASGVTEKRIIDPSLFS 108
DB 502 NINLRPKQNDTKSMAQFCSDTVTTIMHYHGCLVGVKVSPPNRKVLGVDRLRVIDGSTFD 561
QY 109 SGPVTD 114
DB 562 ESRGTN 567
```

RESULT 5

```
US-10-449-902-42587
; Sequence 42587, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences
; APPLICANT: Bio-oriented Technology Research Advancement Institution
; APPLICANT: The Institute of Physical and Chemical Research
; APPLICANT: Foundation for Advancement of International Science
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USBS THEREOF
; FILE REFERENCE: MOA-A0205YI-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 36791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42587
; LENGTH: 794
; TYPE: prt
; ORGANISM: Oryza sativa
US-10-449-902-42587
```

Query Match
8.7%; Score 86; DB 6; Length 794;


```

; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USBS THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53325
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-53325

Query Match
      8.5%; Score 84; DB 6; Length 800;
Best Local Similarity 23.6%; Pred. No. 5.7;
Matches 42; Conservative 26; Mismatches 72; Indels 38; Gaps 7;

QY 12 NSLFGNIGKQSCGTSTASSPCIFRRYPVDCYARAHKMQILMNGDCEK-----QPV 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 215 SSIFFSETTDFNTSMCSBP-----SCSRSLSSSIDLHGNGHVEKISRVSFN 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 66 YGNLKAISTGT-----CCVAMS-YHVAIIVSYKNAAGVTEKRIIDPSLFSGSPVTDI 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 266 LRQIKGLSGSMPLPEPRLRCPISLQMLHDPVIT-----ASQOTERACIEKWFSSGNTTGP 321
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 116 AMRNACVNTS-----CGSASVSSYANTAGNVYRVSNS-----LYDNNLINTNCVLT 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 322 KTRNELSOLSMTPNYCIKGLIASMCQNGVLVPSAPPDPKLYRISLSINSKCLVT 379
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-10-530-983-21
; Sequence 21, Application US/10530983
; Publication No. US20060172301A1
; GENERAL INFORMATION:
; APPLICANT: Innovations Foundation
; TITLE OF INVENTION: Secreted Acid Phosphatase (sapM) is Present Only in Pathogenic
; TITLE OF INVENTION: Mycobacteria and Expressed Selectively at Phagosomal pH
; FILE REFERENCE: 4146 0005
; CURRENT APPLICATION NUMBER: US/10/530,983
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: US 60/416,957
; PRIOR FILING DATE: 2002-10-09
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-530-983-21

Query Match
      8.1%; Score 80.5; DB 6; Length 447;
Best Local Similarity 23.1%; Pred. No. 6.2;
Matches 42; Conservative 25; Mismatches 80; Indels 35; Gaps 9;

QY 33 ITFRYVVDGCVAAHMKROIIMNNGYDCEK-----FVGNLKAISTGTCCVA-W 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 244 ILTFPEDGTVSSKNKIFSVLGGALPDELKGTQDDFTYTHYSVIASVSANWGLPSLGRW 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 81 SYHVALIVSYKNSG-----VTEKRIID--PSLFSGSPVT--DTAMRNACVNTSCGSA 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 304 DCGANILEIYANTGYVNVVDFTNLRINFTYRGPMSAGYSKXSVWPMALLRGDCSAG 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 S-----VSSYANTAGNVYRVSNSLYDN-NLINTNCVLYFSLIS--GCSPPAPDV 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 364 HGILDIVKETYANTETPNYSSP---FPYDTASNVTKVATKKNVTGTRHSSSSSSPSA 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 181 SS 182
   |||
DB 421 SS 422

RESULT 11
US-11-056-355B-97301
; Sequence 97301, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 97301
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(455)
; OTHER INFORMATION: Ceres Seq. ID no. 12736955
US-11-056-355B-97301

Query Match
      8.1%; Score 80.5; DB 7; Length 455;
Best Local Similarity 29.1%; Pred. No. 6.4;
Matches 37; Conservative 15; Mismatches 46; Indels 29; Gaps 8;

QY 60 CEKQFVYGNLKAISTGTC--CVAMS-YHVAIIVSYKNAAGVTEKRIIDPSLFSGSPVTDI 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 CMLQSLIDPTKYTLGLCHBSQPMSTPONTLTILKXKSLCLE-----STGP----- 378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 118 RNACV--NTSCGASVSSYAN-TGANNYY--RSPNSLYLD---NNLINTNCVLTFRS 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 379 -NAPVXLSETSCSSPNLSMEWTISASNNMLAKSTNNSICLDVDETNNLMASNCXKVG-G 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 168 LLSGSCP 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 437 EDSCDP 443
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-11-056-355B-97300
; Sequence 97300, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 97300
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(457)
; OTHER INFORMATION: Ceres Seq. ID no. 12736954
US-11-056-355B-97300
```

```
Query Match      8.1%; Score 80.5; DB 7; Length 457;
Best Local Similarity 29.1%; Pred. No. 6.4;
Matches 37; Conservative 15; Mismatches 46; Indels 29; Gaps 8;

QY 60 CEKQFYGNLKAISTGTC--CVAMSYHVALIVSYKNASGVTEKRIIDPSLFSGGPVTDTAM 117
DB 336 CMLOSILDPKTVLGLCNESQPSWSTYPQNTLTLLKDKSLCLE-----STGP----- 380
QY 118 RNACV---NTSCGSASVSSYAN--TAGNVYY--RSPNSYLYD---NNLINTNCVLTFRS 167
DB 381 -NAPVALSTSCSPMLSEMETISASNMILAAKSTNNSLCDVDETNNLMASCKCVK-G 438

QY 168 LLSGCSGP 174
DB 439 EDSSCDP 445

RESULT 13
US-11-056-355B-97299
; Sequence 97299, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 97299
; LENGTH: 508
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(508)
; OTHER INFORMATION: Ceres Seq. ID no. 12736953
US-11-056-355B-97299

Query Match      8.1%; Score 80.5; DB 7; Length 508;
Best Local Similarity 29.1%; Pred. No. 7.3;
Matches 37; Conservative 15; Mismatches 46; Indels 29; Gaps 8;

QY 60 CEKQFYGNLKAISTGTC--CVAMSYHVALIVSYKNASGVTEKRIIDPSLFSGGPVTDTAM 117
DB 387 CMLOSILDPKTVLGLCNESQPSWSTYPQNTLTLLKDKSLCLE-----STGP----- 431
QY 118 RNACV---NTSCGSASVSSYAN--TAGNVYY--RSPNSYLYD---NNLINTNCVLTFRS 167
DB 432 -NAPVALSTSCSPMLSEMETISASNMILAAKSTNNSLCDVDETNNLMASCKCVK-G 489

QY 168 LLSGCSGP 174
DB 490 EDSSCDP 496

RESULT 14
US-11-174-307B-1872
; Sequence 1872, Application US/11174307B
; Publication No. US20060143729A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 2750-1601PUS2
; CURRENT APPLICATION NUMBER: US/11/174,307B
; PRIOR FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583,671
; PRIOR FILING DATE: 2004-06-30
```

```
; PRIOR APPLICATION NUMBER: 60/583,781
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,651
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 5544
; SEQ ID NO 1872
; LENGTH: 1182
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Pfam Name: peroxidase; Pfam Description: Peroxidase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 21593467; NR Description: peroxidase ATP13a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 31745133; NR Description: putative peroxidase
; OTHER INFORMATION: [Arabidopsis thaliana] >gi|9759059|dbj|BA809581.1| peroxidase
; OTHER INFORMATION: [Arabidopsis thaliana] >gi|1429219|emb|CAA67312.1| peroxidase
; OTHER INFORMATION: ATP13a [Arabidopsis thaliana] >gi|1402900|emb|CAA6966.1|
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 21593692; NR Description: putative peroxidase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 6091756; NR Description: putative peroxidase
; OTHER INFORMATION: [Arabidopsis thaliana] >gi|25453221|sp|O9S867|PE28 APATH Peroxidase
; OTHER INFORMATION: precursor (atperox p28) (ATPp39) >gi|15228606|ref|NP_187017.1|
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 51970002; NR Description: putative peroxidase
US-11-174-307B-1872

Query Match      8.1%; Score 80; DB 7; Length 1182;
Best Local Similarity 20.7%; Pred. No. 22;
Matches 34; Conservative 14; Mismatches 62; Indels 54; Gaps 7;

QY 23 CGTS--TASSPCTITRYVDCGYARAHKMRQILMANGDCEKQFYGNLKAISTGTC-CVA 79
DB 566 CGTTCCTTGTGTC-----CAATA-----TCATATTGGATCATGTCACAG 604
QY 80 WSYHVALIVSYKNASGVTEKRIIDPSLFSGGPVTDTAMRNACVNTSCGSASVSYANTAG 139
DB 605 CTCCTCTTGTGAAGA-----TGCTTTTTA--TGATCTCATCATGATGCTG 652
QY 140 NVYRSPNSYLYDNNLINTNCVLTFRSLLSGCSGPSAPDVSSC 183
DB 653 -----TTGCTCTTGCGGCTGCATCAATAAAC 678

RESULT 15
US-11-289-102-290
; Sequence 290, Application US/11289102
; Publication No. US20060121511A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Hyerin
; APPLICANT: Shaw, Peter M.
; APPLICANT: Clark, Edwin
; TITLE OF INVENTION: BIOMARKERS AND METHODS FOR DETERMINING SENSITIVITY TO
; FILE REFERENCE: 10338 NP
; CURRENT APPLICATION NUMBER: US/11/289,102
; PRIOR FILING DATE: 2005-11-29
; PRIOR APPLICATION NUMBER: US 60/631,993
; PRIOR FILING DATE: 2004-11-30
; NUMBER OF SEQ ID NOS: 395
```

```

; SOFTWARE: Patentin version 3.3
; SEQ ID NO 290
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-289-102-290

```

```

Query Match      8.0%; Score 79.5; DB 7; Length 723;
Best Local Similarity 22.0%; Pred. No. 14;
Matches 35; Conservative 26; Mismatches 71; Indels 27; Gaps 5;

```

```

QY      40  DGCYARAHMROILMNNNGYDCEKOFVYGNLKASTGTCCVAMSYYHVALIVSYKXNAGVTEK 99
      | |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB      218  DACFSKFEHTNNLTWNCENC-GSYCY----SSGPGCQSQKXVFSSTSVTAHKQNSAQIP 272
      | |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      100  RIIDPSLFSSGPTDT-----AMRNACVNTSCGSASVSYANTAGNVY 142
      | |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB      273  VALGKSLRPSAEMIEFTNDSGKTELFCSINCLSAAYRVKTVTSSGVQVVSCHSCKTSAIPQY 332
      | |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      143  YRSPSNSLYDNNLINTNCVL---TKESLLSGCSPSPAP 178
      | |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB      333  HLAMSGTY--SFCSSCVVAFQNVPSKPKGTNSSAVP 369
      | |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

```

Search completed: September 16, 2006, 06:39:38
Job time : 33 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 06:18:35 ; Search time 40 Seconds
(without alignments)
445.003 Million cell updates/sec

Title: US-10-815-774-6

Perfect score: 991
Sequence: 1 LASVIPDVATLNSLFNQIKN.....FSLISGCSPPAPDVSSCGF 185

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96.5	9.7	534	2 C69146	hypothetical prote
2	92	9.3	551	2 A83232	conserved hypothet
3	92	9.3	2395	1 S50820	surface protein ly
4	91	9.2	594	2 T50765	adhesion of calyx
5	90	9.1	243	2 T28802	hypothetical prote
6	89.5	9.0	927	2 T21772	hypothetical prote
7	89	9.0	1895	2 T15881	hypothetical prote
8	88.5	8.9	283	2 G69212	conserved hypothet
9	88.5	8.9	1251	2 T21389	hypothetical prote
10	85.5	8.6	534	2 T39903	serine-rich protei
11	85.5	8.6	1082	2 S37837	suppressor protein
12	85	8.6	594	2 T50764	adhesion of calyx
13	83.5	8.4	712	2 T33028	hypothetical prote
14	82.5	8.3	5376	2 T42215	zonadhesin - mouse
15	82	8.3	703	2 S67277	probable membrane
16	82	8.3	928	2 B72077	polymorphic membra
17	82	8.3	928	2 E86546	polymorphic outer
18	81.5	8.2	361	2 T19647	hypothetical prote
19	81	8.2	485	2 T49237	hypothetical prote
20	79.5	8.0	207	2 T22436	hypothetical prote
21	79.5	8.0	1894	2 UC4980	plexin 1 precursor
22	79	8.0	221	2 A59040	cysteine proteinas
23	79	8.0	275	2 T43004	hypothetical prote
24	79	8.0	732	2 B69749	hypothetical prote
25	78.5	7.9	725	2 T36712	band-6-protein - h
26	78.5	7.9	1067	2 T33174	probable zinc prot
27	77.5	7.8	422	2 T39306	mitogen-activated
28	77	7.8	814	2 T02195	hypothetical prote
29	77	7.8	1237	2 D71850	probable outer mem

30	77	7.8	2090	2 T30075	hypothetical prote
31	77	7.8	2153	2 T30074	hypothetical prote
32	77	7.8	2718	2 A23475	G surface protein
33	76.5	7.7	499	2 S39113	alpha-N-arabinofur
34	76.5	7.7	505	2 T28859	inward rectifier P
35	76.5	7.7	1234	2 C84465	hypothetical prote
36	76	7.7	735	2 A69146	hypothetical prote
37	76	7.7	934	2 T02742	probable ligand-ga
38	76	7.7	2704	2 S09118	G surface protein
39	75.5	7.6	594	2 T31744	hypothetical prote
40	75.5	7.6	1154	2 A86318	protein P15H18.11
41	75	7.6	443	2 F96990	NADP-specific glut
42	74.5	7.5	439	2 A36385	surface antigen se
43	74.5	7.5	980	2 T30089	probable zinc prot
44	74.5	7.5	1046	2 A26838	prestalk protein P
45	74.5	7.5	1335	2 T17508	glycoprotein Vp260

ALIGNMENTS

RESULT 1
C69146 hypothetical protein MTH359 - Methanobacterium thermoautotrophicum (strain Delta H)
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: C69146
R/Smith, D.R.; Doucette-Stamm, L.A.; DeLonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Oliv, D.; Spadefora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Ukwani, N
Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A/Reference number: A69000; MUID:98037514; PMID:9371463
A/Accession: C69146
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residue: 1-534 <MTH>
A/Cross-references: UNIPROT:026459; UNIPARC:UPI000006648A; GB:AE000821; GB:AE000666; NII
A/Experimental source: strain Delta H
C/genetics:
A/Genes: MTH359

Query Match 9.7%; Score 96.5; DB 2; Length 534;
Best Local Similarity 24.1%; Pred. No. 0.75;
Matches 39; Conservative 23; Mismatches 71; Indels 29; Gaps 4;

QY	5	IPDVATLNSLFNQIKNS-----CGTSTASSPCITFRYPVD-----GCY 43
DB	358	LPNTVTVTFLASLKNRPENDPYRGSTARYLASSACPDSPDIRSLASEITRGITSTF 417
QY	44	AAAHK-----MRQILMNGDCEKQFYVGNLKASTGCCVAMSYHVAIVSYKASGVTEK 99
DB	418	SRBAVFGWNRNINVSFYNTKYGAVGTLKRTGCV---DHTHLVALRAAKGIPAR 473
QY	100	RIIDPSLFSSGCVPTDTAMRNACVNTSCGSASVSYANTAGNV 141
DB	474	VYHGTCNFTSSGNVYGHVMAQLLVDPITWYAADATSSNSLSGV 515

RESULT 2
A83232 conserved hypothetical protein PA3310 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: A83232
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Larbig, K.; Lam,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: A83232
A/Status: preliminary

QY 124 TSCGSASVSYANTAGNVTYRSPNSY-LYDNNL-INTNCV 162
DB 143 GVC---VAQGTSYNNAMYSSTGNYCTYDNTCCGANGCV 179

RESULT 6

121772
hypothetical protein T01D3.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21772; T24296
R:Steward, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19470
A:Accession: T21772
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-927 <M12>
A:Cross-references: UNIPROT:O02364; UNIPARC:UPI0000086677; EMBL:Z81527; PIDN:CAB04279.1;
A:Experimental source: clone F35E12
R:Steward, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19870
A:Accession: T24296
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-927 <M12>
A:Cross-references: UNIPARC:UPI0000086677; EMBL:Z81110; PIDN:CAB03262.1; GSPDB:GN00023;
C:Genetics:
A:Gene: CESP:T01D3.6
A:Map position: 5
A:introns: 38/2; 87/3; 128/2; 242/2; 374/3; 416/3; 734/3; 842/3; 857/1

Query Match 9.0%; Score 89.5; DB 2; Length 927;
Best Local Similarity 20.1%; Pred. No. 5.9;
Matches 61; Conservative 25; Mismatches 68; Indels 149; Gaps 16;
QY 4 VIPDVATLNLNFNOI-KNOSGCTASSPCTFRYPVGVGCVARAHKMQI-----LMN 55
DB 130 VEPDVASGFGCQNEIKCKDNKNSGKNA-----DCYVANHQLNCLCKPGYTARR 176
QY 56 NGYDECKQ-----FVYGNLIKAS-TGTCCVAMS----- 81
DB 177 NGRDDCMKQQAQCMGSDPHYVYVYDGRPYQGTCPVFSQPCPTLTPAPYLVMSVAKNEL 236
QY 82 ----YHVA---ILVSYKN-----ASGVTEKRIIDPSLFS-----S 109
DB 237 PGKGYHISQVSEVEVDLNLTHVDGRSKTALVNGV---QVLTPLYFPKNTWTVRVRS 293
QY 110 GP-----VTDLTARNACV-----NTSCGSA-----SVSSYANTAGNV 141
DB 294 GSTFTTENDQGVVVFITTYNSLCVOVPDIPERNGATTLGGLAGNIDGKKLDVYVKNQSV 353
QY 142 Y-----YRSPSN-----SYLYDNNLI-----NTNCVLTRESLFS 170
DB 354 LAIKSRQENNNHADMFKETEDTWTITDKFLIIRPGQENCINGQTLDDNNTNCVSTISLAQ 413
QY 171 GCS 173
DB 414 SCA 416

RESULT 7

T15881
hypothetical protein D1044.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15881
R:Pauley, A.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of C. elegans cosmid D1044.
A:Reference number: Z18423

A:Accession: T15881
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1895 <PAU>
A:Cross-references: UNIPARC:UPI000004CAB8; EMBL:U00065; NID:g495681; PID:g495684; PIDN:
C:Genetics:
A:Gene: CESP:D1044.3
A:introns: 35/2; 63/2; 150/1; 191/3; 207/2; 405/3; 551/3; 629/1; 764/1; 933/1; 959/2; 10

Query Match 9.0%; Score 89; DB 2; Length 1895;
Best Local Similarity 24.1%; Pred. No. 15;
Matches 46; Conservative 27; Mismatches 88; Indels 30; Gaps 9;

QY 20 NOSC-----GTSYASPCITFRYPVGVGCVARAHKMQI-----LMN 55
DB 1112 NQCVYISGLNQCGLTGVSYNSQCITLASFGONCCOSSCINDSVCMQKCNNNRYLV 1171
QY 67 GNLSKSTGTCVAMSRYH---AIVSYKNASGVTEKRIIDPSLFSSGPT---DTAMRN 119
DB 1172 GCVCPITSSICQQTOTLVNQCCLVLSIGETICIANQCVCGAMCNSGTCQCTNGATAMYG 1231
QY 120 ACVNTSCGSASVSYANTAGNVYR-----SPNSYLYDNNLINTN---CVLTRESLISGC 172
DB 1232 YCISSSSSSCN-SNQVSYNGCYTNTVQVGSQCSFSGQCLNNAVCTNNICVSTFCV--SC 1288
QY 173 SPSPAPDVSSC 183
DB 1289 STNQVCISNQC 1299

RESULT 8

G69212
conserved hypothetical protein MTH845 - Methanobacterium thermoautotrophicum (strain Del
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: G69212
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
J.; Olin, D.; Spadafora, R.; Vicaire, R.; Mang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
J.; Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: G69212
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-283 <MTH>
A:Cross-references: UNIPROT:O26933; UNIPARC:UPI000006662A; GB:AE000861; GB:AE000666; NID
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH845

Query Match 8.9%; Score 88.5; DB 2; Length 283;
Best Local Similarity 24.6%; Pred. No. 1.9;
Matches 29; Conservative 17; Mismatches 61; Indels 11; Gaps 2;

QY 41 GCVARAHKMQIILMNNNGYDCEKQFYGNLTKASTGT-----CCVAMSRYHVAIVSYKNA 93
DB 161 GONSTYDKAVRIFFKWTDRSDISYFYNTRKAGAVGTLHSGSANC---DHTHLVALAKRA 216
QY 94 SGVTEKRIIDPSLFSSGPTVTDARNACVNTSCGSASVSYANTAGNVYRSPNSYL 151
DB 217 SGIPARVHMGCVFRSGNTYGHVWQGLYVNGRWYDADATPSNMLGTVMNDRSSAFI 274

RESULT 9

T21389
hypothetical protein F26C11.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21389
R:Matthews, P.
submitted to the EMBL Data Library, December 1994

A:Reference number: Z19416
 A:Accession: T21389
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1251 <WII>
 A:Cross-references: UNIPROT:009550; UNIPARC:UPI0000013BF4C; EMBL:247072; PIRN:CAA87369.1
 A:Experimental source: clone F26C11
 C:Genetics:
 A:Gene: CESP:F26C11.3
 A:Map position: 2
 A:introns: 24/1; 111/1; 208/1; 272/1; 380/1; 394/3; 485/3; 586/3; 630/3; 669/3; 713/3; 7

Query Match
 Best Local Similarity 23.7%; Score 88.5; DB 2; Length 1251;
 Matches 36; Conservative 24; Mismatches 55; Indels 37; Gaps 6;

QY 21 QSC-GTSTASPCITFRYPVDCGYARAHKROQILMNGDYCEKQFYVGNLAKSTGTCCVA 79
 DB 986 QGCKATSTYQTP-TTNNWPTGGT-TTLPDSGEIILSES----- 1021
 QY 80 WSYHVALIVSYKNAAGVTEKRIIDPSLFSSGPVTDAMNACVNTSCGSASVSYANTAG 139
 DB 1022 -----LIAYKCTIVLMQILINPSKNTTRET-TSDAGCKATSSGTTSTWSPGTTGG 1073
 QY 140 NYYRSPSNLYDNNLINTNCVLTKEFLSG 171
 DB 1074 TTVSRFTNSNNPIDSSTLET---TTFAMPPTG 1101

RESULT 10
 T39903
 A:Sequence: rich protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T39903
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
 Submitted to the EMBL Data Library, November 1998
 A:Reference number: Z21889
 A:Accession: T39903
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-534 <LYN>
 A:Cross-references: UNIPROT:O94317; UNIPARC:UPI0000069AEB; EMBL:AL033534; PIRN:CAA22127.
 A:Experimental source: strain 972h-; cosmid c215
 C:Genetics:
 A:Gene: SPDB:SPBC215.13
 A:Map position: 2
 C:Superfamily: pig submaxillary mucin

Query Match
 Best Local Similarity 24.2%; Score 85.5; DB 2; Length 534;
 Matches 45; Conservative 32; Mismatches 86; Indels 23; Gaps 5;

QY 6 PDVATLNLFNQIKQSCGTSTASPCITFRYPVDCGYARAHKROQILMNGDYCEKQFY 65
 DB 78 PTSSSEPTIFSSATPS-ETMSYSSPVSSYSDPATS-----QLPSSSTSPSSSE 127
 QY 66 YGNLKASTGTCCVAMSYHVALIVSYKNAAGVTEKRIIDPSLFSSGPVTDAMNACVNTS 125
 DB 128 YTPSSSTESSLLDPSSVSAILPS-----STVYEVISSSSLSSSDPLTSTSPSSLSSTS 183
 QY 126 CGSASVSS-YANTAGNVYRSPNSGYLYDNNLI-----NTNVCVLTKEFLSGCSPSP 176
 DB 184 SSQPSVSTSTSTSTSSAPFTSTSSYLSSSSSVSSSSSPSSSSSTLTSSSLSTST 243
 QY 177 APDVSS 182
 DB 244 SSSSSS 249

RESULT 11
 S37837
 A:Reference number: SPT23 - yeast (Saccharomyces cerevisiae)
 A:Accession: S37837
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1082 <RIB>
 A:Cross-references: UNIPROT:P35210; UNIPARC:UPI0000135D64; EMBL:Z28020; NID:9486010; PID
 A:Experimental source: strain S288C
 R:Burkett, T.J.; Garfinkel, D.J.
 Yeast 10, 81-92, 1994
 A:Title: Molecular characterization of the SPT23 gene: a dosage-dependent suppressor of
 A:Reference number: S40695; MUID:94262317; PMID:8203154
 A:Accession: S40696
 A:Molecule type: DNA
 A:Residues: 136-714, 'P', 716-737, 'ERKTYLVLLATYGY' <BUR>
 A:Cross-references: UNIPARC:UPI0000168E4A; EMBL:L24760; NID:9404815; PIRN:AAA20575.1; PI
 A:Map position: 111
 F:742-774/Domain: ankyrin repeat homology <AN1>

Query Match
 Best Local Similarity 21.7%; Score 85.5; DB 2; Length 1082;
 Matches 40; Conservative 27; Mismatches 54; Indels 63; Gaps 8;

QY 27 TASSPCITFRYPVD-GCYARAHKROQILMNGDYCEKQFYVGNLAKSTGTCCVAMSYHVA 85
 DB 259 SGNNSCINFDLPTRIVCYCRHKK-----ATNGF----- 288
 QY 86 ILVSYKNAAGVTEKRIIDPSLF-----SSGPVTDAMNACVNTSCGSASVSYANTAGN 140
 DB 289 IFLRDHMDIILAKITTPIMIMDKKASNTTPTSTISMAQVSM--TNDTRSPSSPOSD 346
 QY 141 VYRYS-----PSNSYLYDNNLINTNCVLT-----TKFSLISGCSPPSPAP 178
 DB 347 LMFPEFPPLPSVS---KNFVISTNCMLDSNCGNNNNNDNDNKNNTKITNYAMNNRHFPS 403

QY 179 DVSS 182
 DB 404 NSSS 407

RESULT 12
 T50764
 A:Sequence: adhesion of calyx edges protein ACE (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 05-Oct-2004
 C:Accession: T50764
 R:Araki, T.; Nakatani-Goto, M.
 Submitted to the EMBL Data Library, May 1999
 A:Description: Arabidopsis ADHESION OF CALYX EDGES (ACE), genomic.
 A:Reference number: Z25223
 A:Accession: T50764
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-594 <ARA>
 A:Cross-references: UNIPROT:Q9SX23; UNIPARC:UPI000009C6A5; EMBL:AB027507; PIRN:BAA77842
 A:Experimental source: strain Nossen
 C:Genetics:
 A:Gene: ACE
 A:introns: 25/1; 155/2; 259/1; 387/3; 577/2
 C:Superfamily: alcohol oxidase

Query Match
 Best Local Similarity 24.6%; Score 85; DB 2; Length 594;
 Matches 31; Conservative 13; Mismatches 46; Indels 36; Gaps 5;

QY 25 TSTASPCITFRY---PVD--GCYARAHKROQILMNGDYCEKQFYVGNLAKST-- 73

Db 442 TIVDDNPSTVFYFKHVPDLQRCVEARLVLVKVTSKRFPLNTYQCCKQVNHKMLSLSYVA 501
 QY 74 -----GTCVAMSYVALIVS-----YKASGYTEKRIIDPSLPS 108.
 Db 502 NITLRKQALNDTSMQAFCKDVTVTIWHHGCLVGKVVSPNRKVLGVDRLRVIDDSTDP 561
 QY 109 SGPVTD 114
 Db 562 ESPGTN 567

RESULT 13

T33028
 hypothetical protein T09A12.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T33028
 R:Hawkins, U.; Fullon, B.; Gilling, B.
 submitted to the EMBL Data Library, February 1998
 A:Description: The sequence of C. elegans cosmid T09A12.
 A:Reference number: Z21265
 A:Accession: T33028
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-712 <HAM>
 A:Cross-references: UNIPARC:UPI000017B882; EMBL:AF047660; PIDN:AAC04434.1; GSPDB:GN00022
 A:Experimental source: strain Bristol N2; clone T09A12
 C:Genetic: GSP:T09A12.4
 A:Gene: CBSP
 A:Map position: 4
 A:Introns: 32/1; 158/1; 220/1; 244/3; 295/1; 319/1; 436/3; 543/2; 634/3

Query Match: 8.4%; Score 83.5; DB 2; Length 712;
 Best Local Similarity 21.5%; Pred. No. 15;
 Matches 49; Conservative 20; Mismatches 66; Indels 93; Gaps 11;

QY 23 CGTSTASSPCTTFRYVVDGCVY-RAHKMRQILMNNGYDEKQFVGNLK-ASTGTCVAVW 80
 Db 251 CGTDSNG-----IHFGVDACACSAFFRRIVLVANKDYSCKGKGVVDGSAQKCRAC 305
 QY 81 SYHVALIVSY-----KNASGYTEKRIIDPSL-----FSSGP 111
 Db 306 RFRKCISSGMDKNSVOHRRDAICGYAGV--KRELPPDAEFEPBAKYSTVSEPTSSGS 363
 QY 112 -----VTDAMRNACVNTSGSASVSS-----YANT 137
 Db 364 GGNQNVSSPAGIPRVPTLRTTQASTCMMSCGQKSVLHELICRONFLTEOROLFYAGC 423
 QY 138 AGNVYRSPNSNYLDNNLINTNCVLTFSLSGCSPPAPDVSSCGF 185
 Db 424 LGD-WERKPS--IENQTL-----SELTDFSSCMF 450

RESULT 14

T42215
 zonadhesin - mouse
 N:Alternate names: sperm-specific membrane protein
 C:Species: Mus musculus (house mouse)
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C/Accession: T42215
 R:Guo, Z.; Garbers, D.L.
 J. Biol. Chem. 273, 3415-3421, 1998
 A:Title: Species diversity in the structure of zonadhesin, a sperm-specific membrane protein
 A:Reference number: Z22080; PMID:9813114; PMID:9452463
 A:Accession: T42215
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-5376 <GAO>
 A:Cross-references: UNIPROT:O88799; UNIPARC:UPI000002A15F; EMBL:U97068; NID:G3327420; PI
 C:Genetic:
 A:Gene: Zan
 A:Map position: 5

C:Function:
 A:Description: functions in multiple cell adhesion processes
 A>Note: found exclusively on the apical region of the sperm head
 C:Keywords: cell adhesion

Query Match: 8.3%; Score 82.5; DB 2; Length 5376;
 Best Local Similarity 24.9%; Pred. No. 1.8e+02;
 Matches 43; Conservative 12; Mismatches 53; Indels 65; Gaps 10;

QY 24 GTS-TASSPCTTFRYVVDGCVYRAHKMRQILMNNGYDEKQFVGNLKASTGTCVAVW 82
 Db 4530 GTSFAPSTC-----EEBCV-----CEPDVYLSNDKCVPSSEC----- 4562
 QY 83 HVALIVSYKASGYTEKRIIDPSLPSGPVTDVW--RNACVNTSGSASVSSYANTAGN 140
 Db 4563 -----GCKDAIGV---LIPES-----KTVWSRCKTNCCTCKGQGVQCHD----- 4598
 QY 141 VYRSPNSNYLDNNLINTNCV-----LTKFSLSGCSPPAPDVSS-----CG 184
 Db 4599 --FSCPTGSRCDNNEGNSNCVTYALKCPAHLTYNCLPSCLPSCSDPBLGCG 4649

RESULT 15

S67277
 probable membrane protein YOR365C - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein O6661
 C:Species: Saccharomyces cerevisiae
 C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
 C/Accession: S67277
 R:Delius, H.; Hebling, U.; Hofmann, B.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67261
 A:Accession: S67277
 A:Molecule type: DNA

A:Residues: 1-703
 A:Cross-references: UNIPROT:Q08844; UNIPARC:UPI000006A32E; EMBL:Z75272; NID:G1420789; PI
 A:Experimental source: strain S288C
 C:Genetic:
 A:Gene: MIPS:YOR365C
 A:Cross-references: SGD:S0005892
 A:Map position: 15R
 A:Superfamily: probable membrane protein YAL053W
 C:Keywords: transmembrane protein
 F:2-18/Domain: transmembrane #status predicted <TM1>
 F:208-224/Domain: transmembrane #status predicted <TM2>
 F:248-264/Domain: transmembrane #status predicted <TM3>
 F:379-395/Domain: transmembrane #status predicted <TM4>
 F:431-447/Domain: transmembrane #status predicted <TM5>
 F:454-470/Domain: transmembrane #status predicted <TM6>
 F:546-562/Domain: transmembrane #status predicted <TM7>
 F:575-591/Domain: transmembrane #status predicted <TM8>
 F:599-615/Domain: transmembrane #status predicted <TM9>
 F:616-632/Domain: transmembrane #status predicted <TM10>

Query Match: 8.3%; Score 82; DB 2; Length 703;
 Best Local Similarity 19.0%; Pred. No. 20;
 Matches 39; Conservative 34; Mismatches 66; Indels 66; Gaps 6;

QY 31 PCITRRYP-----VDGCVYRAHKMRQILMNNGYDEKQFVGN----- 68
 Db 157 PSITYNIPDLDAIVYVTAYSTIDKEFTPLACVQWLSNGRVTQREYLSWNLVILITSGI 216
 QY 69 -----LKASNGTCVAMSYVALIVSYKN-----ASGV 96
 Db 217 MGSVYSLQGYVTTRLASISISLLYFONLAILAMISVSLPPIVAANTONFQWSMGI 276
 QY 97 TE-----KRIIDPSLPSGPVTDVWNRACV-----NTSGASVSSYANTAGNYRSP 146
 Db 277 IININMQRFLDMVYVAATSSPTVYVRNKEVLSISVQKSLNSKIISASNLNG--IBSRO 334
 QY 147 SNSIYLDNNLINTNCVLTFSLSG 171
 Db 335 KNDLLYTSNLRNSNDYLSKIIVLRG 359

Mon Sep 18 16:01:20 2006

us-10-815-774-6.rpr

Page 6

Search completed: September 16, 2006, 06:24:04
JOB time : 43 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 06:15:05 ; Search time 298 Seconds

(without alignments)
574.255 Million cell updates/sec

Title: US-10-815-774-6

Perfect score: 991
Sequence: 1 LASVIPDVATLNSLFNQIKN.....FSLSGCSPSPADVSSCGF 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	320	Q9A008_9FLAO	Q9A08 chryseobact
2	142.5	14.4	252	Q6MGS8_BDEBA	Q6MGS8 bdellovibri
3	127.5	12.9	274	Q5W27_LEGRL	Q5W27 legionella
4	127.5	12.9	294	Q5ZUX4_LEGRL	Q5ZUX4 legionella
5	125.5	12.7	274	Q5X4N7_LEGRL	Q5X4N7 legionella
6	111.5	11.3	267	Q6MJD7_BDEBA	Q6MJD7 bdellovibri
7	96.5	9.7	534	Q26459_METHA	Q26459 methanobact
8	93.5	9.4	572	Q70R15_GIALA	Q70R15 giardia lam
9	92	9.3	551	Q9HYT4_PSEAE	Q9HYT4 pseudomonas
10	92	9.3	2385	Q27167_PARTE	Q27167 parametium
11	91	9.2	294	Q56X03_ARATH	Q56X03 arabidopsis
12	91	9.2	594	HTH_ARATH	HTH_ARATH
13	90	9.1	243	Q18804_CAEEL	Q18804 caenorhabdi
14	90	9.1	405	Q7YX35_CRYPV	Q7YX35 cryptospori
15	89.5	9.0	927	Q7JXS6_CAEEL	Q7JXS6 caenorhabdi
16	89	9.0	672	PON25_CAEEL	P41951 caenorhabdi
17	89	9.0	956	Q60MG7_CAEEL	Q60MG7 caenorhabdi
18	88.5	8.9	283	Q26933_METHA	Q26933 methanobact
19	88.5	8.9	1317	YQUS_CAEEL	Q09550 caenorhabdi
20	88.5	8.9	1572	Q61UF2_CAEEL	Q61UF2 caenorhabdi
21	88	8.9	267	Q7R89_PSEAE	Q7R89 pseudomonas
22	88	8.9	419	Q86FM2_CLOTE	Q86FM2 clostridium
23	88	8.9	457	Q55FE6_DICDI	Q55FE6 dictyosteli
24	87.5	8.8	406	Q5CMJ9_CRYPV	Q5CMJ9 cryptospori
25	87.5	8.8	1506	Q5CMK7_CRYPV	Q5CMK7 cryptospori
26	86.5	8.7	1155	Q54NC7_DICDI	Q54NC7 dictyosteli
27	86	8.7	344	CYSP5_DICDI	P54640 dictyosteli
28	86	8.7	344	Q8MNB1_DICDI	Q8MNB1 dictyosteli
29	86	8.7	529	Q8DSM3_VIBRU	Q8DSM3 vibrio vuln
30	86	8.7	789	Q48BL4_COLEP3	Q48BL4 colwellia p
31	86	8.7	794	Q769C9_ORYSA	Q769C9 oryza sativ

32	86	8.7	794	2	Q76C22_ORYSA	Q76C22 oryza sativ
33	86	8.7	794	2	Q76C26_ORYSA	Q76C26 oryza sativ
34	86	8.7	1219	2	Q8W3E4_ORYSA	Q8W3E4 oryza sativ
35	86	8.7	2397	2	Q607Y4_PABPR	Q607Y4 parametium
36	85.5	8.6	534	2	Q94317_SCHPO	Q94317 schizosacch
37	85.5	8.6	1082	1	SP723_YEAST	P35210 saccharomyc
38	85.5	8.6	1213	2	Q4DD54_TREYR	Q4DD54 trypanosoma
39	85.5	8.6	1213	2	Q81E41_PLAF7	Q81E41 plasmodium
40	85	8.6	684	2	Q4SX0_TETNG	Q4SX0 tetradodon n
41	85	8.6	958	2	Q7QUB2_GIALA	Q7QUB2 giardia lam
42	85	8.6	1420	2	Q45YJ8_SCHMD	Q45YJ8 schmidtea m
43	85	8.6	1594	2	Q81ZN7_PLAF7	Q81ZN7 plasmodium
44	84.5	8.5	240	2	Q61H40_CAEEL	Q61H40 caenorhabdi
45	84.5	8.5	719	2	Q54192_DICDI	Q54192 dictyosteli

ALIGNMENTS

RESULT 1

ID	Q9A008_9FLAO	PRELIMINARY;	PRT;	320 AA.
AC	Q9A008;			
DT	01-JUN-2001,	integrated into UniProtKB/TrEMBL.		
DT	01-JUN-2001,	sequence version 1.		
DT	07-FEB-2006,	entry version 10.		
DE	Protein-glutaminase precursor.			
GN	Name=prgA,			
OS	Chryseobacterium proteolyticum.			
OC	Bacteria; Bacteroidetes; Flavobacteriales;			
OC	Flavobacteriaceae; Chryseobacterium.			
OX	NCBI_TaxId=118127;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=9670;			
RA	MEDLINE=21153247; PubMed=11231294;			
RT	Yamaguchi S., Jeenes D.J., Archer D.B.;			
RT	"Protein-glutaminase from Chryseobacterium proteolyticum, an enzyme			
RT	that deamidates glutamyl residues in proteins: purification,			
RT	characterization and gene cloning";			
RL	Eur. J. Biochem. 268:1410-1421(2001).			
CC				
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms			
CC	Distributed under the Creative Commons Attribution-NonDerivs License			
CC	EMBL: AB046594; BAB21508.1; -; Genomic_DNA.			
KW	Signal.			
FT	SIGNAL 1 21			
FT	CHAIN 136 320			
SQ	SEQUENCE 320 AA; 35044 MW; C67823D2BC131410 CRC64;			

Query Match 100.0%; Score 991; DB 2; Length 320;
Best Local Similarity 100.0%; Pred. No. 6.4e-86;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1 LASVIPDVATLNSLFNQIKNOSCGSTASSPCTFRYPVPGCGYAAHKKROILMANGYDC 60
DB	136 LASVIPDVATLNSLFNQIKNOSCGSTASSPCTFRYPVPGCGYAAHKKROILMANGYDC 195
QY	61 EKQFYGNLKAISTGTCVAVSYHVALIVSYKASGVTEKRIIDPSLFSSGPVTDIARNA 120
DB	196 EKQFYGNLKAISTGTCVAVSYHVALIVSYKASGVTEKRIIDPSLFSSGPVTDIARNA 255
QY	121 CVNTSCGASVSSYANTAGNYRRPSNSYLYDNNLITNCLTKFSLSGCSPSPADPV 180
DB	256 CVNTSCGASVSSYANTAGNYRRPSNSYLYDNNLITNCLTKFSLSGCSPSPADPV 315
QY	181 SSCGF 185
DB	316 SSCGF 320

RESULT 2

```
Q6MGS8_BDEBA
ID Q6MGS8_BDEBA PRELIMINARY; PRT; 252 AA.
AC Q6MGS8;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DE 07-FEB-2006, entry version 9.
DI Hypothetical protein precursor.
GN OrderedLocustNames=Bd3844;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;

[1]
NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendall S., Jędrzej P., Rosinski A., Eppinger M., Baar C.,
RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,
RA Socke R.E., Schuster S.C.,
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective."
RL Science 303:689-692(2004).

-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
-----
EMBL; BX842656; CAB81201.1; -; Genomic_DNA.
DR BioCyc; BBAC264462; BD3844-MONOMER; -.
KW Complete proteome; Hypothetical protein; Signal.
FT STGNL 1 Potential.
SQ SEQUENCE 252 AA; 26633 MW; 8B80A8661F93118C CRC64;

Query Match 14.4%; Score 142.5; DB 2; Length 252;
Best Local Similarity 32.3%; Pred. No. 4e-05;
Matches 40; Conservative 17; Mismatches 54; Indels 13; Gaps 4;

QY 5 IPDVATNSLFNQK-----NOSCGTSVASSPCITFRYPVGCYARAHKMQIIMNGYDC 60
DB 66 IPDVGSYADLEREKYYVDTRFTADAPFARRLTWMYTPDCCARALATLVEHQFTT 125
QY 61 EKO-FVYGNLKAATGTC---CVAMSYHVALIVSKYKASGVTEKRIIDPSLFSSGPTDTA 116
DB 126 PKKIFVGNLYAPFQNAFGSVSWMYHVAIVYRVGN-----EYVFPDAINPEKPTLVE 180
QY 117 WRNA 120
DB 181 WNKA 184

RESULT 3
ID Q5WM27_LEGPH PRELIMINARY; PRT; 274 AA.
AC Q5WM27;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DE 07-FEB-2006, entry version 8.
DI Hypothetical protein.
GN OrderedLocustNames=lp11633;
OS Legionella pneumophila (strain Lens).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=297245;

[1]
NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC PubMed=15467720; DOI=10.1038/ng1447;
RA Cazale C., Rusniok C., Brueggemann H., Zidane N., Magnier A., Ma L.,
RA Tichit M., Jarraud S., Bouchier C., Vandenesch F., Kunst F.,
RA Etienne J., Glaser P., Buchrieser C.,
RT "Evidence in the Legionella pneumophila genome for exploitation of
RT host cell functions and high genome plasticity."
RL Nat. Genet. 36:1165-1173(2004).

-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
```

```
CC Distributed under the Creative Commons Attribution-NonDerivs License
-----
EMBL; CR628337; CAH15873.1; -; Genomic_DNA.
DR LegiList; lp11633; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 274 AA; 31408 MW; 2E87A5626A211293 CRC64;

Query Match 12.9%; Score 127.5; DB 2; Length 274;
Best Local Similarity 31.9%; Pred. No. 0.0012;
Matches 44; Conservative 23; Mismatches 35; Indels 37; Gaps 8;

QY 3 SVIPDVAT--LNSLFNQIKNSCGTSTASP-----CITFRYPVDCYARAHKMQI 52
DB 72 STVPKVSYSYELAMFNILRD---TRFLYSPDPDPQRRISWLYPDDCFARALSRIK 127
QY 53 LMNNGYDCE-----KQFVGNLK-----ASTGTCVAMSYHVALIVSKYKASGVTEKRI 101
DB 128 L-----DSEHFVIPAKIFVGDLEMQTPYSSEGT--VSMWYHVASVANNYKGTI-----YV 175
QY 102 IDPSLFSSGPTDTAMRN 119
DB 176 LDPAPKBPGRPMLIDWVN 193
```

```
RESULT 4
ID Q5ZUX4_LEGPH PRELIMINARY; PRT; 294 AA.
AC Q5ZUX4;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DE 07-FEB-2006, entry version 7.
DI Hypothetical protein.
GN OrderedLocustNames=lp91668;
OS Legionella pneumophila subsp. pneumophila (strain Philadelphia 1 /
OS ATCC 33152).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=272624;

[1]
NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC PubMed=15448271; DOI=10.1126/science.1099776;
RA Chien M., Morozova I., Shi S., Sheng H., Chen J., Gomez S.M.,
RA Asanani G., Hill K., Nuara J., Feder M., Rineer J., Greenberg J.J.,
RA Stesenko V., Park S.H., Zhao B., Teplickaya E., Edwards J.R.,
RA Pamou S., Georgiou A., Chou I.-C., Iannuccilli W., Uiz M.E.,
RA Kim D.H., Geringer-Sameth A., Goldsberry C., Morozov P., Fischer S.G.,
RA Segal G., Qu X., Rzhetsky A., Zhang P., Cayonis E., De Jong P.J.,
RA Du J., Kalachikov S., Shuman H.A., Russo J.J.,
RT "The genomic sequence of the accidental pathogen Legionella
RT pneumophila."
RL Science 305:11966-1968(2004).

-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
-----
EMBL; AB017354; AAU27748.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 294 AA; 33731 MW; 8850B3A8D18669CD CRC64;

Query Match 12.9%; Score 127.5; DB 2; Length 294;
Best Local Similarity 31.2%; Pred. No. 0.0013;
Matches 43; Conservative 23; Mismatches 35; Indels 37; Gaps 8;

QY 3 SVIPDVAT--LNSLFNQIKNSCGTSTASP-----CITFRYPVDCYARAHKMQI 52
DB 92 STVPKVSYSYELAMFNILRD---TRFLYSPDPDPQRRISWLYPDDCFARALSRIK 147
QY 53 LMNNGYDCE-----KQFVGNLK-----ASTGTCVAMSYHVALIVSKYKASGVTEKRI 101
DB 148 L-----DSEHFVIPAKIFVGDLEMQTPYSSEGT--VSMWYHVASVANNYKGTI-----YV 195
QY 102 IDPSLFSSGPTDTAMRN 119
```

Db 196 LDPAAKEGPMILDDWYN 213

RESULT 5
OSX4N7_LEGPA PRELIMINARY; PRT; 274 AA.
ID 05X4N7_LEGPA
AC 05X4N7;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Hypothetical protein.
GN Ordered locus names=Jp01639;
OS Legionella pneumophila (strain Paris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_Taxid=297246;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15467720; DOI=10.1038/ng1447;
RA Cazalet C., Rusniok C., Brueggemann H., Zidane N., Magnier A., Ma L.,
RA Ticht M., Jarrad S., Bouchier C., Vandenesch F., Kunst F.,
RA Etienne J., Glaser P., Buchrieser C.;
RT "Evidence in the Legionella pneumophila genome for exploitation of
host cell functions and high genome plasticity."
RT Nat. Genet. 36:1165-1173(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; CR628336; CAH12791.1; -; Genomic DNA.
DR LegioLst; Jp1639; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 274 AA; 31284 MW; 5B302361FELC834B CRC64;

Query Match 12.7%; Score 125.5; DB 2; Length 274;
Best Local Similarity 31.2%; Pred. No. 0.0018;
Matches 43; Conservative 22; Mismatches 36; Indels 37; Gaps 8;

QY 3 SVTPDVAT---LNSLPNOKKSCGTSASP-----CTFRYPVDCVYARAHMKRI 52
DB 72 STVPKVSSTYDELNNMNLIRD---TRFLYSPDKPFRIRISLVPDDCFAPAAALSRK 127
53 LNNNGYDCE-----KQFYGNLK-----ASTGTCVANSYHVALIVSYKNSGVTEKRI 101
DB 128 L-----DSHFVTPAKITFAFGDLEMOTPTSSSEST--VSMWYHVSANVNNMGTI-----YV 175
QY 102 IPPSLFSSGCVPTDTAMRN 119
DB 176 LDPAAKEGPMILDDWYN 193

RESULT 6
Q6MJD7_BDEBA PRELIMINARY; PRT; 267 AA.
ID 06MJD7;
AC 06MJD7;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Hypothetical protein.
GN Ordered locus names=Bd2842;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_Taxid=959;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RC PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,
RA Sockett R.B., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
genomic perspective.";

RL Science 303:689-692(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; BX842653; CAE80623.1; -; Genomic DNA.
DR BioCyc; BBAC264462; BD2842-MONOMER; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 267 AA; 29959 MW; 9AD3258D47989064 CRC64;

Query Match 11.3%; Score 111.5; DB 2; Length 267;
Best Local Similarity 26.3%; Pred. No. 0.038;
Matches 45; Conservative 28; Mismatches 63; Indels 35; Gaps 9;

QY 2 ASVIDPVATLNS-----LFNQIKKSCGTSASPCTFRYPVDCVYARAHMKRIIDPSL 55
DB 68 ATLQDVKATVNSSEADAKLFLSKSQG-----DIPFERSLNGCERRAHMSRLML 118
QY 56 NGYDEKQFVYGN-----LKASTGT---CCVANSYHVALIVSYKNSGVTEKRIIDPSL 106
DB 119 KGITPLKVPASVNEDESRLRPNTKNGMTYDMKHYAPVLYKKGSILV--PYWMDPSL 177
QY 107 FSSG-PVTDAMR-----NACVNTSCGASVSYANTAGNVYRSPNSY 150
DB 178 EKKAVPVS--WQATMTRHNPQKVKYLKFTPATY--NDAGITRVNFKONDF 225

RESULT 7
O26459_METTH PRELIMINARY; PRT; 534 AA.
ID 026459_METTH
AC 026459;
DT 01-JAN-1998, integrated into UniProtKB/TrEMBL.
DT 01-JAN-1998, sequence version 1.
DT 07-FEB-2006, entry version 23.
DE Conserved protein.
GN Ordered locus names=MT359; ORF names=MT3_359;
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanobacter.
OX NCBI_Taxid=187420;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=Delta H;
RC MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Delouchey C., Lee H.-M., Dubois J.,
RA Altredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lunn W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics."
RT J. Bacteriol. 179:7135-7155(1997).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AE000666; AAB84865.1; -; Genomic DNA.
DR PIR; C69146;
DR BioCyc; MTH359-MONOMER; -;
DR InterPro; IPR002931; Tmsglumase_like.
DR pfam; PF01841; Tmsglum_core; 1.
DR SMART; SM00460; TGC; 1.
KW Complete proteome.
SQ SEQUENCE 534 AA; 57628 MW; 588412BD8D342E26 CRC64;

Query Match 9.7%; Score 96.5; DB 2; Length 534;
Best Local Similarity 24.1%; Pred. No. 2.4; Indels 29; Gaps 4;
Matches 39; Conservative 23; Mismatches 71;

QY 5 IPDVATLNSLPNOKKSCGTSASPCTFRYPVDCVYARAHMKRIIDPSL 52
DB 72 STVPKVSSTYDELNNMNLIRD---TRFLYSPDKPFRIRISLVPDDCFAPAAALSRK 127

```

DB      358 LPNTVTRILASLKNRPENDPYRGESTARIASSASCPVDSPIRSLJSEITRGLTSTF 417
OY      44 ARAHK-----MQIILMNGYDCEKQFVGNLKAISTGCCVANSYVAIIVSYKMSGVTEK 99
DB      418 SRAEAVFGVRDININVSFFYNTKYAGVGLNKRITNCV-----DTHHLVALARAAGLPAR 473
OY      100 RIIDBSLFSSGPVTDAMRNACVNTSCGSASVSANTAGNY 141
DB      474 VYHGTCNFTSGNVYGHVMAQLVGDTPWYAADATSSRSLGVV 515

RESULT 8
OY      07ORIS_GIALA PRELIMINARY; PRT; 572 AA.
ID      07ORIS_GIALA
AC      07ORIS_GIALA
DT      15-DEC-2003, integrated into UniProtKB/TrEMBL.
DT      15-DEC-2003, sequence version 1.
DE      07-FEB-2006, entry version 15.
DE      GUP 576.21850.23568.
OS      Giardia lamblia ATCC 50803.
OC      Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
OX      NCBI_TaxID=184922;
RN      NUCLEOTIDE SEQUENCE.
RP      STRAIN=WB C6;
RC      Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA      Olsen G.J., Sogin M.L.;
RT      "Draft sequence of the Giardia lamblia genome."
RL      Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC      -!- CAUTION: The sequence shown here is derived from an
          EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
          preliminary data.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NonDerivs license
CC      -----
DR      EMBL; AACB01000139; EAA37471.1; -; Genomic_DNA.
DR      HSSP; P16581; 1G1T.
DR      InterPro; IPR006210; EGF.
DR      InterPro; IPR000742; EGF_3.
DR      InterPro; IPR013111; EGF_extrecell.
DR      InterPro; IPR013032; EGF_like_reg.
DR      Pfam; PF01683; EB; 2.
DR      SMART; SM00181; EGF; 5.
DR      PROSITE; PS00022; EGF_1; UNKNOWN_4.
DR      PROSITE; PS01186; EGF_2; 5.
DR      PROSITE; PS00026; EGF_3; 3.
DR      SEQUENCE 572 AA; 61787 MW; 92B8AF72A01C17BB CRC64;
SQ

Query Match          9.4%; Score 93.5; DB 2; Length 572;
Best Local Similarity 22.2%; Pred. No. 4.9;
Matches 45; Conservative 22; Mismatches 57; Indels 79; Gaps 10;

OY      32 CTFRRPVVGCYARAH--GMRQILMNGYD-----CEKQFVGNLKAISTGTC--- 76
DB      210 CWSAENPNVFCSHGHCDETRKICIDGIDGVCQKYKPEEGDGVY-----SBQCCKD 264
OY      77 -CVANSYVAIIVSYKMSGVTEKRIIDS--LFSSGPVTD----- 114
DB      265 TCVAQDGHVCA----KHGSKTESCVCDPQVYLIGAECTPAVCLVGEVCPHGECVFM 320
OY      115 -----TAMRAACVNTSCGSAS-----VSSYANTA-----GNVYY 143
DB      321 NOSYCRCKDAEYTAFFENKCIPINSCISATFPYGVPELCSNKGTCDDMKRRCVCPPLVGGTYC 380
OY      144 RSPSNSYL--YDNNLINTNCVLTT 164
DB      381 QKCEDALAMEDSCVSLNCIST 403

RESULT 9
OY      09HT4_PSEAE PRELIMINARY; PRT; 551 AA.
ID      09HT4_PSEAE

```

```

AC      09HT4;
DT      01-MAR-2001, integrated into UniProtKB/TrEMBL.
DT      01-MAR-2001, sequence version 1.
DT      07-FEB-2006, entry version 19.
DE      Hypothetical protein.
CN      OryedredlocusNames=PA3310;
OS      Pseudomonas aeruginosa.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Pseudomonas.
OX      NCBI_TaxID=287;
RN      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP      STRAIN=ATCC 15692 / PAO1.
RC      MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RK      Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA      Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagron M.,
RA      Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA      Brody L.L., Coulter S.N., Folger K.R., Kas A., Latbig K., Lim R.M.,
RA      Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA      Reizer J., Sailer M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT      "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
          opportunistic pathogen."
RL      Nature 406:959-964(2000).
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NonDerivs license
CC      -----
DR      EMBL; AE004753; AAC06698.1; -; Genomic_DNA.
DR      PIR; A83232; A83232.
DR      BioCyc; PAER287; PA3310-MONOMER; -.
DR      GO; GO:0008484; F:sulfuric ester hydrolase activity; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR012549; DUF1705.
DR      InterPro; IPR000917; Sulfatase.
DR      Pfam; PF08019; DUF1705; 1.
DR      Pfam; PF00884; Sulfatase; 1.
KW      Complete proteome; Hypothetical protein.
SQ      SEQUENCE 551 AA; 61422 MW; 353EAF5C308BE815 CRC64;

Query Match          9.3%; Score 92; DB 2; Length 551;
Best Local Similarity 21.6%; Pred. No. 6.6;
Matches 42; Conservative 24; Mismatches 54; Indels 74; Gaps 9;

OY      15 FNOIKNKGSGSTA--SSPCTFRPVVGC---YARAK-----MQIILMNGYDC 60
DB      274 FTQVR--SGTSTAVSVPMSCMSQYPEREDYSPKAKTHEGLDITIQKAGVOVLMENNSDC 331
OY      61 EKQFVGNLKAISTGTC-----CVANSYVAIIVSYKMSGVTEK 99
DB      332 K-----GTLKVPNRDIPKTPSPPCDGNKLDLSLVGL---QEYIDGLQDD 376
OY      100 RII-----DPSLFSSGPVTDAMRNACVNTSCGSASVSANTAGNYRRSPNSYL 151
DB      377 AIIVLHSDSHGPEYERYPERKEMERFPQVCRITNOLGSCSKBELVN-----V 422
OY      152 YDNNLINTNCVLTK 165
DB      423 YDNTIIVTDHFLTK 436

RESULT 10
OY      027167_PARTE PRELIMINARY; PRT; 2395 AA.
ID      027167_PARTE
AC      027167_PARTE
DT      01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT      01-NOV-1996, sequence version 1.
DT      07-FEB-2006, entry version 27.
DE      51B type surface protein (Surface antigen 51B).
OS      Parametium tetraurelia.
OC      Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
OC      Parametidae; Parametium.
OX      NCBI_TaxID=5888;
RN      [1]

```


RESULT 13
 Q18804 CAEEL PRELIMINARY; PRT; 243 AA.
 ID Q18804 CAEEL
 AC Q18804
 DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
 DT 01-NOV-1996, sequence version 1.
 DE Hypothetical protein C53B7.3.
 GN ORFNames=C53B7.3;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_Taxid=6239;
 RX MEDLINE=99069613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL, U42830; AAC48274.1; -; Genomic_DNA.
 DR PIR, T28802; T28802.
 DR HSRP, P02876; 9MGA.
 DR Ensembl, C53B7.3; Caenorhabditis elegans.
 DR Wormbase, WBGene00016894; C53B7.3.
 DR WormPep, C53B7.3; CE06973.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 243 AA; 26299 MW; 6C1E33CB1CA2D16B CRC64;
 SQ
 Query Match 9.1%; Score 90; DB 2; Length 243;
 Best Local Similarity 26.7%; Pred. No. 3.9;
 Matches 43; Conservative 14; Mismatches 64; Indels 40; Gaps 9;
 QY 22 SGGTSSASPCITFRPVGCVARAHKMKQILMNNGYDEKQFVGNLKAQT--GTCCVA 79
 DB 39 SGGTSSASPCITFRPVGCVARAHKMKQILMNNGYDEKQFVGNLKAQT--GTCCVA 90
 QY 80 WSHVAIVSYKASGVTEKRIIDPSLFS-----SGPV-----TDTAMRACVN 123
 DB 91 -----TSYSCRSEFTCMNRCCSSYNNMYNSQYSNGSGSTRYSTDCTSGCMCSN 142
 QY 124 TSCGSASVSSYANTAGNVYRSPSNYS-LYDNNL-INTNCV 162
 DB 143 GVC---VAQYGTSYNNAMYSSTGNYCTYDNTCGINQRCV 179
 RESULT 14
 Q7Y235_CRYPV PRELIMINARY; PRT; 405 AA.
 ID Q7Y235_CRYPV
 AC Q7Y235;
 DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2003, sequence version 1.
 DE Hypothetical protein.
 DT 07-FEB-2006, entry version 7.
 GN ORFNames=IMB.718;
 OS Cryptosporidium parvum.
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 OC Cryptosporidiidae; Cryptosporidium.
 NCBI_Taxid=5807;
 RX NUCLEOTIDE SEQUENCE.
 RA Bankier A.T., Spriggs H.F., Partmann B., Konfortov B.A., Madera M.,
 RA Vogel C., Teichmann S.A., Ivens A., Dear P.H.;
 RT "Integrated mapping, chromosomal sequencing and sequence analysis of
 Cryptosporidium parvum."
 RL Genome Res. 0:0-0(2003).
 CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL, BX538353; CAD98294.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 405 AA; 46655 MW; 4FC406FA937929EF CRC64;
 Query Match 9.1%; Score 90; DB 2; Length 405;
 Best Local Similarity 28.8%; Pred. No. 7.1;
 Matches 32; Conservative 18; Mismatches 47; Indels 14; Gaps 5;
 QY 66 YGNLKAATGTCVANSYVAIVSYKASGVTEKRIIDPSLFSGGPVDTAMRACVNTS 125
 DB 45 YGSLH-----FHRENFRCRIKEKGLSGLEKSLID-FIDPLSHLDKGTITLNNHS 96
 QY 126 CGSASVSSYANTAGNVYRSPSNYSY-----YDNNLINTN-C-VLTKFSILS 170
 DB 97 DFGDLPQOMQUTGSGVYTRGDSQVLLAQAIDSSIRNPECEVTTEFALIA 147
 RESULT 15
 Q7YK56 CAEEL PRELIMINARY; PRT; 927 AA.
 ID Q7YK56 CAEEL
 AC Q7YK56;
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DE Hypothetical protein.
 GN ORFNames=T01D3.6; T01D3.6a;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_Taxid=6239;
 RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RG MEDLINE=99069613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL, Z81527; CAB04279.1; -; Genomic_DNA.
 DR EMBL, Z81110; CAB04279.1; JOINED; Genomic_DNA.
 DR EMBL, Z81527; CAB03262.1; -; Genomic_DNA.
 DR EMBL, Z81527; CAB03262.1; JOINED; Genomic_DNA.
 DR Wormbase, WBGene00011330; T01D3.6.
 DR WormPep, T01D3.6a; CE12964.
 DR GO, GO:0005509; F:calcium ion binding; IEA.
 DR InterPro, IPR000152; Asx_hydroxyl_S.
 DR InterPro, IPR006210; EGF.
 DR InterPro, IPR000742; EGF 3.
 DR InterPro, IPR001881; EGF_Ca_bd.
 DR InterPro, IPR013091; EGF_Ca_bd_2.
 DR InterPro, IPR013032; EGF_like_reg.
 DR InterPro, IPR002181; Fibrinogen_C.
 DR InterPro, IPR002919; Prot_inh_CR_TTL.
 DR InterPro, IPR001846; VWF D.
 DR Pfam, PF07645; EGF CA; 1.
 DR Pfam, PF00147; Fibrinogen_C; 1.
 DR Pfam, PF01826; TTL; 1.
 DR Pfam, PF00094; VMD; 1.
 DR SMART, SM00181; EGF 3.
 DR SMART, SM00179; EGF CA; 1.
 DR SMART, SM00216; VEGF 1.
 DR SMART, SM00216; VEGF 1.
 DR PROSITE, PS00010; ASX_HYDROXYL; 1.
 DR PROSITE, PS00022; EGF 1; UNKNOWN_2.
 DR PROSITE, PS01186; EGF_2; 2.

DR PROSITE: PS50026; EGF 3; 3.
 DR PROSITE: PS01187; EGF_CA; 1.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 927 AA; 103414 MW; 71AEF0A61FC2B266 CRC64;

Query Match 9.0%; Score 89.5; DB 2; Length 927;

Best Local Similarity 20.1%; Pred. No. 21; Mismatches 68; Indels 149; Gaps 16;
 Matches 61; Conservative 25;

```

QY 4 VLPDVATLNSLFEVQI--KNQSCGISTASSPICITFRYPVDGCTARAHKMQI-----LMN 55
DB 130 VCPDYASGRFCQNEIKCKXNSGKNA-----DCYVANHQLNCICKPGYTARR 176
QY 56 NGYDCEKQ-----FVYGNLKAS-TGTCVAVNS-----S 81
DB 177 NGRDCDMKVQOACMSGDPHYVTYDGLRFYQGTCPYVFSQPCTTLPAPYLMYSVARAKNEL 236
QY 82 ----YHVA----ILVSYKN-----ASGVTEKRIIDPSLFS-----S 109
DB 237 PKGYHISQVSEVEVDLNLITIHVDGRSKTALVNGV--QVLTPWYFPNKNTWTVVRVRS 293
QY 110 GP-----VTDTAMFNACY-----NTSCGSA-----SVSSYANTAGNV 141
DB 294 GSTFTIENDQGVVVTFTYNSLCVQVPDIPEFNGATTLGSLAGNIDGKKLDVNVKNGSV 353
QY 142 Y----YRSPSN-----SYLYDNNLI-----NTNCVLTKEPSLIS 170
DB 354 LAIKSSKROPENNNHADFMKTEDTWTIDKFLILRPQENGINQOTLDNNTNCVSTSIISLAQ 413
QY 171 GCS 173
DB 414 SCA 416
```

Search completed: September 16, 2006, 06:23:20
 Job time : 302 secs